Gravity as a diffeomorphism-invariant gauge theory

Kirill Krasnov

School of Mathematical Sciences, University of Nottingham, Nottingham, UK
(Received 23 March 2011; published 19 July 2011)

A general diffeomorphism-invariant SU(2) gauge theory is a gravity theory with two propagating polarizations of the graviton. We develop this description of gravity, in particular, for future applications to the perturbative quantization. Thus, the linearized theory, gauge symmetries, and gauge fixing are discussed in detail, and the propagator is obtained. The propagator takes a simple form of that of Yang-Mills theory with an additional projector on diffeomorphism-equivalence classes of connections inserted. In our approach the gravitational perturbation theory takes a rather unusual form in that the Planck length is no longer fundamental.

I. INTRODUCTION

The field theory approach to gravity, see, e.g., [1], tells us that gravity is not a gauge theory. Indeed, the carriers of force in a gauge theory (such as, e.g., Maxwell electrodynamics) are spin-one particles. For this reason there are two types of charged objects interacting by exchange of carriers of force—those negatively and positively charged. Like particles repel and unlikes attract. In contrast, there is only one type of charge in gravity and everything attracts everything. Thus, gravity is not a gauge theory; see [1] for a more detailed discussion.

This simple argument forbids a direct gauge theory description of gravity. It says nothing, however, about less direct possible relationships. And indeed, a relation of a completely different type is now becoming very popular. This has its origin in the open-closed string duality, which implies that amplitudes for closed strings are squares of those for open strings. Since the low-energy limit of the closed string theory is gravity, and that for open strings is gauge theory, this implies that scattering amplitudes for gravitons must be expressible as squares of amplitudes for gluons, see, e.g., a review [2] and/or a more recent paper [3] and references therein. The relationship is not direct, and it is, in particular, not easy to find a Lagrangian version of the correspondence. However, it has recently led to some very interesting developments on loop divergences in $N = 8$ supergravity. Another example of a gauge theory/gravity relation is the AdS/CFT correspondence [4] of string theory.

The aim of the present paper is to develop (further, see historical remarks below) yet another gravity/gauge theory correspondence. Currently there appears to be no relation between the present story and that of [2]. The relationship of interest for us here has its origins in the discovery of Plebanski [5] that a certain triple of self-dual two-forms can be used as the basic variable for gravity.[1] The same

"self-dual" formulation of general relativity (GR) has been rediscovered a decade later by Ashtekar [7] via a completely different path of a canonical transformation on the phase space of GR. The two discoveries were later linked in [8], and the outcome was a realization that gravity can be reformulated as a theory whose phase space coincides with that of an SU(2) gauge theory. This gravity/gauge theory relationship was taken one step further in [9]. Thus, it was realized that the two-form fields of Plebanski formulation of GR [5] can be integrated out to obtain a "pure connection" formulation of general relativity, where the only dynamical field is an SU(2) connection. The result was a completely new perspective on general relativity, in which GR becomes reformulated as a novel type of theory of the gauge field—a *diffeomorphism-invariant gauge theory*.

The work on the pure connection formulation of GR [9] has led to some further advances in that it was realized in [10] that there is not a single diffeomorphism-invariant gauge theory, but an infinite-parameter class of them. All these theories share the same key properties with GR, as they have the same number of propagating degrees of freedom (DOF). Thus, for any theory in the class the phase space is that of an SU(2) gauge theory. However, in addition to the usual SU(2) gauge rotations, there are also diffeomorphisms acting on the phase space variables, which reduce the number of propagating DOF from 6 of SU(2) gauge theory to 2 of GR.

Unfortunately, the new pure connection viewpoint on GR originating in [9] (and having its roots in Plebanski's key insight [5]) has not been significantly developed. The phase space version [7] of this story has formed the foundation of the approach of loop quantum gravity [11], but the pure connection formulation of GR [9] and of the infinite-parameter family [10] of "neighbors of GR" has not had any significant applications, as far as the author is aware. The main aim of this paper is to revisit this pure connection formalism for gravity and develop it further. Our main motivation is a (future) application of this formalism to the perturbative quantization of gravity. However, as will become clear below, the pure connection perspective on gravity developed here may have other uses.

---

[1]Similar ideas have appeared in the literature much earlier, see, e.g., [6] for historical remarks, but it was Plebanski who proposed reformulating general relativity without the metric, with only two-forms as dynamical variables.

We motivate our interest in this formalism for (quantum) gravity with some historical remarks. Thus, the author's interest in the subject started in [12] from a simple power counting argument describing how the nonrenormalizability of GR manifests itself in the Plebanski formulation [5]. The outcome was an infinite-parameter family of Plebanski-like theories, where the constraint term of the Plebanski action was replaced by a "potential" term for the would-be Lagrange multipliers. Each of the new theories is just the familiar from discussions of nonrenormalizability counterterm-corrected GR (in disguise), and so the interpretation of the infinite number of new parameters is that they are related to coefficients in front of counterterms constructed from the curvature and its derivatives in the usual metric description of gravity. It was very quickly realized [13] that the new infinite-parameter family of theories [12] is essentially the same as the one introduced and studied by Begtsson and collaborators a decade earlier [10], with the difference being that [10] worked at the level of a pure connection formulation, while the theories [12] are formulated as Plebanski-like theories with two-form fields as the basic variables.

The class of gravity theories [10,12] can be thought of as summing at least some of the quantum corrections that arise in the process of renormalization of GR, and in [14] this was confirmed by directly exhibiting the familiar GR counterterms as appearing from [12]. The work [12] also conjectured that this class of gravity theories sums up *all* the arising quantum corrections; in other words, it was conjectured that the class [12] is closed under the renormalization, and that the arising renormalization group flow is that in the space of potential functions defining the theory.

At the time of writing [12] the only motivation for this conjecture was the author's optimism—the conjecture did not contradict anything one knew about the nonrenormalizability of GR, and was the most optimistic scenario for how the divergences of GR might organize themselves. The remark [13] relating the Plebanski-like theories [12] to the pure connection theories [10] brought with it an additional justification. Thus, a closer look at these theories made it clear that they are just the most general diffeomorphism-invariant gauge theories. The class of such theories should therefore be closed under the renormalization, because any counterterm that can be needed for canceling the arising quantum divergences is already included into the action; see [15] for the first time this argument was spelled out.

One of the aims of the present paper is to set the stage for a systematic study of the quantum perturbation theory for the gravitational theories introduced in [12] (and previously in [10]). Our final goal is to settle the status of the conjecture of [12] that this class of theories is closed under the renormalization, and then to compute the resulting renormalization group flow. However, it would be impractical to try to write up all the necessary calculations in a single paper. For this reason, in the present paper we develop the classical theory to the extent that the propagating degrees of freedom (gravitons) are manifest. We also do some preliminary steps necessary for the perturbative loop computations in that the gauge fixing is discussed in detail and the propagator is obtained. It is then straightforward to start to compute loop diagrams. This is however not attempted in the present paper, and the task of developing a sufficiently economical way to study the renormalization is left to future work.

Apart from just setting the stage for future quantum calculations, a somewhat unexpected outcome of this work is a completely new viewpoint on the gravitational perturbation theory. As we shall see, in the present diffeomorphism-invariant gauge-theoretic approach to gravity, the fundamental scale is set not by the Newton's constant, which does not appear in the original formulation of the theory at all, but rather by the radius of curvature of the background that is used to expand the theory around. Thus, the natural fundamental length scale is set by the cosmological constant. This has the effect that in our theory the Newton's constant becomes a derived quantity. This leads to some puzzles about the cutoff scale of our perturbation theory, to be discussed towards the end of the paper. Another point that is worth emphasizing from the outset is that our gauge-theoretic approach to gravity only works for a nonzero value of the cosmological constant Λ. As we shall see more explicitly below, the actions we work with blow up in the limit $\Lambda \rightarrow 0$. The puzzles about the behavior of the perturbation theory that we discuss in the main text are directly related to this feature.

To summarize, the main aim of this work is to develop a new approach to the gravitational perturbation theory, for future use, in particular, in the quantum loop calculations. What makes this paper distinct from previous works (in particular of this author) is that here for the first time the pure connection formalism close in spirit to the formulation in [10] is used as a starting point for the gravitational perturbation theory. Thus, all previous works on theories [12] used the two-form formulation. The gravitational perturbation theory in the two-form formulation is similar to that in the usual metric approach, see [14]. In particular, the fundamental scale that determines the self-coupling of the gravitons and sets the scale of the strong coupling regime is, as in the usual metric case, the Planck scale. However, the number of field components one has to work with in the two-form formulation is quite large—it is that of an SU(2) Lie algebra-valued two-form field. Moreover, there are second-class constraints that require the path integral measure to be somewhat nontrivial. For all these reasons it proved to be rather difficult to set up an economical perturbation theory in the two-form formalism. At the same time, for a long time it seemed that the pure connection formulation is ill suited for being a starting point of a perturbative description, as it was not at all clear

how one can expand the theory around the Minkowski spacetime background which corresponds to a zero connection, see, e.g., remarks in [16].

In this paper the prejudices about the pure connection formulation of gravity are put aside and this formulation is used as a starting point for the gravitational perturbation theory. And, as we hope to convince the reader, this formulation can be used rather effectively, in that the arising perturbation theory is reasonably economical. In particular, the linearized theory is very simple (arguably simpler than in the metric description), and the propagator can be obtained without too much difficulty. As we shall see, in the pure connection formalism developed here gravity becomes not too dissimilar to SU(2) gauge theory, the main difference being that a certain additional projector on diffeomorphism equivalence classes is inserted into the standard $1/k^2$ propagator of the gauge theory. This gives hope that the renormalization in this class of gravity theories will eventually become manageable. As we have already mentioned, this is left to the future work.

What is new in this work as compared to previous works on the pure connection formulation, in particular, the work [10] and works by Bengtsson and collaborators that followed, is that our treatment uses in an essential way the formulation in terms of a homogeneous potential function applied to a matrix-valued four-form. This was developed in earlier works of the author, and first spelled out in [17] for the version of the theory that uses a two-form field, and in [15] for the pure connection formulation. This formulation renders the action principle of the theory very compact, and makes it possible to set up the perturbation theory without too much difficulty.

Before we proceed with a description of the theory, there are a few things that ought to be emphasized to avoid misunderstanding. In our gauge-theoretic approach to gravity the theory (or any of the class of theories that we study) remains as nonrenormalizable as GR in the usual metric-based treatment. Thus, as we shall explicitly see below, the coupling constant of our theory has a negative mass dimension, which signals nonrenormalizability by power counting. Thus, the final goal of our enterprise is not to show that the theory is renormalizable—it is not— but rather to show that the infinite-parameter class of theories that we study is closed under the renormalization, and then to compute the arising renormalization group flow. In other words, we are not after the renormalizability in the usual sense of quantum field theory, which is that a Lagrangian with a finite number of couplings is closed under the renormalization. Rather, we are after the renormalizability in the effective field theory sense of Weinberg, see, e.g., [18] for a recent discussion, where any theory is renormalizable once all possible counterterms are added to the action. Our aim is then to show that in the case of gravity in four spacetime dimensions it is sufficient to consider only those counterterms (infinite in number) that

can be compactly summed up into our diffeomorphism-invariant gauge theory Lagrangian. Should this indeed be the case, the renormalization group flow in the infinite dimensional space of gravity theories will be just a flow in the space of defining functions, and will become manageable. Note once again, however, that the quantum theory, while being our main motivation, is not the subject of the present work.

I would also like to explain at the outset how a gauge theory (with spin-one excitations) can describe gravity with its spin-two excitations. This is a version of the story ''spin-one plus spin-one is spin-two'', of relevance for the gauge theory gravity relationship [2]. There are, however, also significant differences. Thus, the main dynamical field of our theories is an SU(2) connection A_μ^i, where μ is a spacetime index, and $i = 1, 2, 3$ is a Lie algebra one. Let us recall that in the usual gauge theories in Minkowski spacetime the temporal component A_0^i of the connection field becomes a Lagrange multiplier—the generator of the gauge rotations. Then of the spatial components A_a^i, where $a = 1, 2, 3$ is a spatial index, some components are pure gauge in that they can be set to any desired value by a gauge transformation. The physical propagating degrees of freedom of the theory can be described as the gauge equivalence classes of the spatial projection of the connection. In the case of gauge group SU(2), the gauge invariance removes three of the nine components of the spatial connection A_a^i, leaving two propagating polarizations per each Lie algebra generator.

As we shall see below, in the case of our gravitational theories the situation is very similar, with the exception that the Lagrangian is in addition invariant under diffeomorphisms. The way this is realized in our theories is that the Lagrangian is simply independent of certain four combinations of the connection field A_μ^i. This is where the spin-two comes from. Thus, consider once again the spatial projection of the connection A_a^i. We shall see that (using the background) it will be possible to identify two types of indices — the spatial and the internal Lie algebra ones. Once this is done, the spatial connection can be thought of as a 3×3 matrix, or, in representation theoretic terms, it constitutes the spin-one tensor spin-one representation. This decomposes as spin-two plus spin-one plus spin-zero. On the other hand, the temporal component of the connection A_0^i forms the-spin one (adjoint) representation of SU(2). The diffeomorphism invariance projects out the spin-zero components of the spatial connection A_a^i, as well as a certain combination of the spin-one component of A_a^i and A_0^i, leaving only one of these spin-one components in the game. Thus, after the projection induced by the diffeomorphisms, the Lagrangian depends only on the spin two component of A_a^i, as well as on the spin-one set of Lagrange multipliers—generators of SU(2) rotations. These make the three longitudinal components of the five-component spin-two field unphysical, leaving only two propagating

physical modes. To summarize, in our version of gauge theory/gravity correspondence the spin-two also comes from the tensor product of two spin-one representations. As in any gauge theory in Minkowski space, one of these spin-one representations is supplied by the spatial projection of the connection field. The other spin-one is provided by the adjoint representation of the SU(2) Lie algebra in which the connection field takes values.

Our final remark here is about the issues of the reality conditions. As we shall see below, in the physically realistic case of Lorentzian signature gravity, the main dynamical field of our theories is a complex-valued $SO(3, \mathbb{C})$ connection. Thus, appropriate reality conditions need to be imposed to select the field configurations corresponding to real Lorentzian metrics. Our strategy for dealing with these in the present paper is as follows. We shall see that at the level of the linearized theory the reality conditions are straightforward (one can easily determine them from the requirement that the linearized Hamiltonian is positive-definite). In the full theory, however, the reality conditions need to be imposed nonperturbatively. We do not yet know how to do this. However, for many applications, in particular, the ones we are most interested in, this is not needed. Thus, for calculations studying the renormalization of gravity (e.g., ones done using the background field method), one performs the Wick rotation to the Riemannian signature metrics. For the later case our gauge field is a real-valued SO(3) connection, and the reality conditions are straightforward. Similarly, for the perturbative loop computations one uses the knowledge of the linearized reality conditions to specify the physical external states. It is then only necessary to specify the contour in the complex connections space that is used in the loop integrations. There is typically very little freedom in the choice of this contour provided one wants the integrals to converge. So, again it is possible to perform computations without specifying explicitly the full nonlinear theory reality conditions. This state of affairs, is, of course, not completely satisfactory, for one would like to have a complete control over the full theory reality conditions as well. This is, however, beyond the scope of this paper (and is never needed here).

With these preparatory remarks having been made, we can proceed to describe how gravity can be reformulated as a diffeomorphism-invariant gauge theory. The organization of the paper is as follows. In Sec. II we define an action principle for our theories, explain how their parametrization by a homogenous function works, derive the field equations and verify gauge invariances of the action. Sec. III studies the theory linearized around a constant curvature background. In particular, a simple quadratic in the gauge field fluctuations action is obtained, and its Hamiltonian analysis is performed. This confirms the outlined above picture of how the spin-two nature of excitations comes about. Sec. IV is central to our analysis. It

discusses the gauge fixing appropriate to the situation at hand, and inverts the gauge-fixed quadratic form to obtain the propagator. In Sec. V we derive the (cubic and quartic) interaction terms of our theory. We conclude with a brief discussion.

II. DIFFEOMORPHISM-INVARIANT GAUGE THEORIES

A. Gravity as a gauge theory

In the pure connection formulation gravity becomes the most general diffeomorphism-invariant gauge theory. In the case of a purely gravitational theory[2] the gauge group is (complexified) SU(2) ∼ SO(3). The action is a functional of an SU(2) connection A^i, $i = 1, 2, 3$ on a spacetime manifold M. Let $F^i = dA^i + (1/2)\epsilon^{ijk}A^j \wedge A^k$ be the curvature of A^i. The action is given by the following gauge and diffeomorphism-invariant functional of the connection:

$$S[A] = (1/\mathrm{i}) \int_M f(F^i \wedge F^j). \tag{1}$$

Here $\mathrm{i} = \sqrt{-1}$ is a factor introduced for future convenience, and f is a function with properties to be spelled out below.

We shall refer to f as the defining function of our theory. It is a holomorphic, homogeneous of degree one and gauge-invariant function of its matrix (and four-form) valued argument. Thus, let $X^{ij} \in \mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$ be a matrix valued in the second symmetric power of the Lie algebra. The gauge group SU(2) ∼ SO(3) acts in the space of such matrices via $X \to gXg^T$, where T is the operation of the transpose. We first consider scalar valued functions $f: \mathfrak{su}(2) \otimes_S \mathfrak{su}(2) \to \mathbb{C}$ that are holomorphic, gauge-invariant $f(gXg^T) = f(X)$ and homogeneous of degree one $f(\alpha X) = \alpha f(X)$. A convenient for practical computations parametrization of such functions is as follows. Consider the following three SU(2) invariants of X^{ij}:

$$\mathrm{Tr}\,(X), \qquad \mathrm{Tr}(X^2), \qquad \mathrm{Tr}(X^3), \tag{2}$$

where the traces (and powers of X) are computed using the Killing metric δ^{ij} on the Lie algebra. When $\mathrm{Tr}(X) \neq 0$ we can parametrize the defining function f as follows:

$$f(X) = \mathrm{Tr}(X)\chi\left(\frac{\mathrm{Tr}(X^2)}{(\mathrm{Tr}(X))^2}, \frac{\mathrm{Tr}(X^3)}{(\mathrm{Tr}(X))^3}\right), \tag{3}$$

where χ is now an arbitrary holomorphic function of its two arguments.

Given f with the properties as spelled out above, e.g., one parametrized as in (3), it can be seen that this function can be applied to a matrix-valued four-form, with the result being a four-form. Indeed, consider $F^i \wedge F^j$, which is a

---

[2]One can also consider unified Yang-Mills-gravity theories of the same sort, see [19].

$\mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$ valued four-form. Choose a reference volume form on M (we assume that M is orientable), and denote it by (vol). Of course, (vol) is only defined modulo the multiplication by a nowhere zero function. Using this reference volume form we can write $F^i \wedge F^j = X^{ij}$(vol), where X^{ij} is again defined only modulo rescalings. We can now use the homogeneity of f to write

$$f(F^i \wedge F^j) = (\text{vol})f(X). \tag{4}$$

It is moreover clear that the result on the right-hand side does not depend on which reference volume form is used in this argument. This is again due to the homogeneity of f. This shows that the integrand in (1) is a well-defined four-form that can be integrated to obtain the action. This finishes the formulation of our theory.

We note that, as formulated, there are no dimensionful parameters in our theory. Indeed, we assume the connection field A^i to have the usual mass dimension one, so that the curvature has the mass dimension two, and the matrix of the wedge products $[X] = 4$. The defining function f is essentially the function χ of ratios of powers of X^{ij} that are dimensionless, and so does not contain any dimensionful parameters (but contains an infinite number of dimensionless ''coupling constants,'' once expanded appropriately). Thus, due to the homogeneity of f, its mass dimension is the same as that of X (in the parametrization (3) the mass dimension is carried by the first term $\text{Tr}(X)$, while the function χ is dimensionless). The function f can then be integrated to produce a dimensionless action (as usual we work in the units $c = \hbar = 1$). As we shall see, the fact that there are no dimensionful coupling constants in our theory has profound implications for the structure of its perturbation theory.

Classically (1) is a theory that can be shown, see e.g. [20], to propagate two (complex for the time being, reality conditions will be discussed below) degrees of freedom. We will see a version of the argument that leads to this conclusion below when we consider the perturbation theory. One can also show that theory (1) is a gravity theory, in spite of the fact that no metric is present anywhere. Thus, it can be reformulated explicitly as a theory of metrics via a sequence of transformations. The main idea is to note that declaring the three two-forms F^i to span the space of (anti-) self-dual two-forms determines a conformal metric on M whenever the matrix X^{ij} defined from the wedge product of curvatures is nondegenerate. One can then rewrite the theory (1) explicitly as the theory of this metric, see [14] for details. We also note that the usual general relativity (with or without the cosmological constant) can be rewritten in this language, see below. However, in this paper, we shall not need this relation to metric theories. Our plan is to study (1) as is. We shall set the stage for its perturbative quantization and a study of its renormalization. The main justification for this undertaking is that a whole class of gravity theories (for varying defining

functions f) can be treated in one go. Moreover, our theories are theories of a connection, and we can hope to use the expertise that was accumulated in quantum field theory for dealing with quantum gauge theories.

B. GR with the cosmological constant

Before we proceed with our analysis of theories (1), we would like to state the action principle that reformulates the usual GR (with the cosmological constant) in this language. Consider the following action principle:

$$S_{\text{GR}}[A] = \frac{1}{16\pi i G\Lambda} \int (\text{Tr}\sqrt{F^i \wedge F^j})^2. \tag{5}$$

Here $\text{Tr}\sqrt{F^i \wedge F^j}$ is the trace of a matrix square root of the matrix $F^i \wedge F^j$. It is clear that the above action is of the general form (1), for the scalar function that is used in the action (5) is gauge-invariant and homogeneous of degree one, as required. The quantities G, Λ in the denominator in front of the action are the usual Newton's constant and the cosmological constant, respectively. Note that these appear in the action only in the dimensionless combination $G\Lambda$. For the currently accepted value of Λ the value of $G\Lambda$ is exceedingly small $G\Lambda \sim M_\Lambda^2/M_p^2 \sim 10^{-120}$. Thus, the value of the dimensionless parameter in front of GR action is very large. Below we shall see what kind of implication this has for the gravity perturbation theory. For the convenience of the reader the action (5) is derived in the Appendix.

C. Topological action

Another prominent member of the class of theories (1) is the topological theory

$$S_{\text{top}}[A] = \frac{1}{i\kappa} \int \text{Tr}(F^i \wedge F^j), \tag{6}$$

where κ is some numerical parameter. It is not hard to see that the Lagrangian is a total derivative, and so the action describes a theory without propagating DOF—a topological theory. Being topological, this theory is certainly a fixed point of the (sought) renormalization group flow in the space of theories (1). Conjecturally, the renormalization group flow takes one from (5) at low energies to (6) at high energies.

D. A convex functional

We would also like to give an example of an action with a convex (near the point $X^{ij} \sim \delta^{ij}$) defining function. Let us consider the following theory:

$$S[A] = \frac{1}{i\tilde{\kappa}} \int \frac{\text{Tr}(F^i \wedge F^j)^2}{\text{Tr}(F^i \wedge F^j)}, \tag{7}$$

where $\tilde{\kappa}$ is a dimensionless parameter. The downward gradient flow for this functional of the connection has been studied in [21]. The renormalization group flow for

the class of theories (1) should, in particular, explain why at low energies one flows to a concave functional (5) instead of a convex functional such as (7).

E. First variation and field equations

The first variation of the action (1) gives us the field equations. To write these down, let us give a parametrization of the matrix X^{ij} useful for practical computations. Thus, let $\tilde{\epsilon}^{\mu\nu\rho\sigma}$ be a completely antisymmetric rank-4 vector, which is a density of weight one (as is indicated by the tilde over its symbol). This object exists on any orientable manifold and does not need a metric for its definition. Consider

$$\tilde{X}^{ij} := \frac{1}{4} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} F^j_{\rho\sigma}, \qquad (8)$$

where as before $F^i_{\mu\nu}$ is the curvature two-form, with its spacetime indices now indicated explicitly. The quantity $\tilde{X}^{ij}$ is a $\mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$ valued matrix, and a density of weight one. One takes the defining function f to be a function of $\tilde{X}^{ij}$ given by the same expression as in (3). With convention $dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma = \tilde{\epsilon}^{\mu\nu\rho\sigma} d^4 x$ we can write the action (1) as

$$S[A] = (1/\mathrm{i}) \int_M d^4 x f(\tilde{X}^{ij}). \qquad (9)$$

The first variation of the action can now be easily computed and reads

$$\delta S[A] = (1/\mathrm{i}) \int_M d^4 x \frac{\partial f}{\partial \tilde{X}^{ij}} \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\rho} \delta A^j_\sigma. \qquad (10)$$

Integrating by parts, we see that the field equations for (1) can be written as

$$D_A B^i = 0, \qquad (11)$$

where we have used the form notations again, and the two-form B^i is defined via

$$B^i := \frac{\partial f}{\partial \tilde{X}^{ij}} F^j. \qquad (12)$$

We note that the matrix of first derivatives that appear on the right-hand side of this expression is a symmetric matrix, and has density weight zero (as the ratio of the density weight one function $f(\tilde{X})$ and the density weight one quantity $\tilde{X}$). Thus, (12) is a well-defined two-form.

For example, in the case of GR we get

$$B^i_{\mathrm{GR}} = \frac{\mathrm{Tr}\sqrt{X}}{16\pi G\Lambda} ((\sqrt{X})^{-1})^{ij} F^j. \qquad (13)$$

Then, using the definition of $X^{ij} \sim F^i \wedge F^j$ we can easily see that in the case of GR

$$B^i_{\mathrm{GR}} \wedge B^j_{\mathrm{GR}} \sim \delta^{ij}, \qquad (14)$$

which is the usual "metricity" equation of Plebanski formulation of GR [5]. Another example is that of the topological theory (6). In this case the B-field is given by

$$B^i_{\mathrm{top}} = \frac{1}{\kappa} F^i, \qquad (15)$$

and the field Eq. (11) is satisfied automatically as a consequence of the Bianchi identity $D_A F = 0$.

F. Symplectic structure

The computation of the first variation in the previous subsection also gives us the symplectic structure of the theory. Thus, the phase space of the theory is the space of all solutions of (11), and the symplectic structure can be obtained by considering the boundary term that was neglected in passing from (10) to (11). The integral of the boundary term gives rise to an integral over the spatial slice Σ of the following quantity

$$\Theta := \frac{1}{2\mathrm{i}} \int_\Sigma B^i \wedge \delta A^i, \qquad (16)$$

where B^i is as in (12). This is a one-form on the phase space of the theory. Its exterior derivative produces the symplectic two-form. We see that the significance of the quantity B^i defined by (12) is that its spatial projection plays the role of the momentum canonically conjugate to the spatial projection of the connection A^i. We emphasize that in the present "pure gauge" formulation, the two-form B^i is not independent and is a function of the connection field. A formulation that "integrates in" the two-form field as an independent variable is possible, and has been studied in previous works by the author, but will not be considered here.

G. Gauge invariance

Let us now verify by an explicit computation that our theory is invariant under diffeomorphisms as well as $SO(3, \mathbb{C})$ rotations. This is of course expected, because the action was constructed in the way that these invariances should hold. However, an explicit verification of this fact will allow us to establish some identities for the latter. The gauge transformations act on the connection field as follows

$$\delta_\xi A^i_\mu = \xi^\alpha F^i_{\mu\alpha}, \qquad \delta_\phi A^i_\mu = D_{A\mu} \phi^i. \qquad (17)$$

The first of these transformations can be seen to be a diffeomorphism corrected by a gauge transformation, while the second one is the usual gauge rotation with the parameter ϕ^i.

It is not too difficult to prove the invariance of our action (1) under these transformations. Let us first consider the diffeomorphisms. The variation of the action (10) becomes proportional to

$$\int_M d^4 x \frac{\partial f}{\partial \tilde{X}^{ij}} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\rho} \xi^\alpha F^j_{\sigma\alpha}. \qquad (18)$$

We now need some identities. First we note that one can write the Bianchi identity $D_A F^i = 0$ as

$$D_{A[\mu}F^i_{\nu]\rho} = -\frac{1}{2}D_{A\rho}F^i_{\mu\nu}. \qquad (19)$$

Another identity that we need is

$$\tilde{\epsilon}^{\,\mu\nu\rho\sigma}F^{(i}_{\mu\nu}F^{j)}_{\sigma\alpha} = -\frac{1}{4}\delta^\rho_\alpha \tilde{\epsilon}^{\mu\nu\gamma\delta}F^i_{\mu\nu}F^j_{\gamma\delta} = -\delta^\rho_\alpha \tilde{X}^{ij}, \quad (20)$$

where δ^ρ_α is the Kronecker delta. Note that the symmetrization is taken on the left-hand side. The above two identities, as well as the definition (8) of the matrix $\tilde{X}^{ij}$, allow us to rewrite (18) as

$$-\int_M d^4x \frac{\partial f}{\partial \tilde{X}^{ij}}\left(\tilde{X}^{ij}\partial_\alpha \xi^\alpha + \frac{1}{2}\tilde{\epsilon}^{\mu\nu\rho\sigma}F^i_{\mu\nu}\xi^\alpha D_{A\alpha}F^j_{\rho\sigma}\right)$$
$$= -\int_M d^4x \frac{\partial f}{\partial \tilde{X}^{ij}}D_{A\alpha}(\xi^\alpha \tilde{X}^{ij}). \qquad (21)$$

Integrating by parts, this becomes equal to

$$\int_M d^4x \xi^\alpha \tilde{X}^{ij}D_{A\alpha}\frac{\partial f}{\partial \tilde{X}^{ij}}. \qquad (22)$$

We should now see that the integrand here is zero. This follows from the homogeneity of the function f. Indeed, we have

$$\tilde{X}^{ij}\frac{\partial f}{\partial \tilde{X}^{ij}} = f \qquad (23)$$

from the fact that f is a homogeneous function of degree one. Let us now apply the operator of partial derivative ∂_μ to both sides of this equation. We get

$$(\partial_\mu \tilde{X}^{ij})\frac{\partial f}{\partial \tilde{X}^{ij}} + \tilde{X}^{ij}\partial_\mu \frac{\partial f}{\partial \tilde{X}^{ij}} = \partial_\mu f = \frac{\partial f}{\partial \tilde{X}^{ij}}\partial_\mu \tilde{X}^{ij}. \quad (24)$$

Comparing the two sides we see that

$$\tilde{X}^{ij}\partial_\mu \frac{\partial f}{\partial \tilde{X}^{ij}} = 0, \qquad (25)$$

which is almost the integrand in (22), except for the fact that we have the covariant derivative in (22). Let us now consider the difference between the covariant and the usual derivatives. We have

$$\tilde{X}^{ij}(D_{A\mu} - \partial_\mu)\frac{\partial f}{\partial \tilde{X}^{ij}} = 2\tilde{X}^{ij}\epsilon^{ikl}A^k_\mu \frac{\partial f}{\partial \tilde{X}^{lj}}. \qquad (26)$$

The quantity here is zero in view of the gauge invariance of the function f. Indeed, under infinitesimal gauge transformations an $\mathfrak{su}(2)\otimes_S\mathfrak{su}(2)$-valued matrix $\tilde{X}^{ij}$ transforms as

$$\delta_\phi \tilde{X}^{ij} = \epsilon^{ikl}\phi^k \tilde{X}^{lj} + \epsilon^{jkl}\phi^k \tilde{X}^{il}. \qquad (27)$$

Then the statement that f is an SO(3, $\mathbb{C}$) invariant function becomes

$$\epsilon^{ikl}\tilde{X}^{kj}\frac{\partial f}{\partial \tilde{X}^{lj}} = 0, \qquad (28)$$

which can be expressed in words by saying that the commutator of the matrix $\tilde{X}^{ij}$ with the matrix $\partial f/\partial \tilde{X}^{ij}$ of the first derivatives of the defining function is zero.

The identity (28) immediately implies that the difference of the derivatives in (26) is zero and thus

$$\tilde{X}^{ij}D_{A\mu}\frac{\partial f}{\partial \tilde{X}^{ij}} = 0, \qquad (29)$$

which proves the invariance of the action (1) under diffeomorphisms.

Let us now prove the invariance of (1) under the gauge rotations. The variation of the action in this case becomes proportional to

$$\int_M d^4x \frac{\partial f}{\partial \tilde{X}^{ij}}\tilde{\epsilon}^{\mu\nu\rho\sigma}F^i_{\mu\nu}D_{A\rho}D_{A\sigma}\phi^j. \qquad (30)$$

Expressing the commutator of the covariant derivatives as the commutator with the curvature, and recalling the definition (8) of the matrix $\tilde{X}^{ij}$ we get

$$4\int_M d^4x \frac{\partial f}{\partial \tilde{X}^{ij}}\epsilon^{jkl}\tilde{X}^{ik}\phi^l, \qquad (31)$$

which is zero in view of (28). This proves the invariance of the action (1) under the SO(3, $\mathbb{C}$) rotations.

H. Second variation

We can now compute the second variation, in preparation for the next section treatment. We have

$$\delta^2 S[A] = (1/\mathrm{i})\int_M d^4x \left(\frac{\partial^2 f}{\partial \tilde{X}^{ij}\partial \tilde{X}^{kl}}\delta \tilde{X}^{ij}\delta \tilde{X}^{kl} + \frac{\partial f}{\partial \tilde{X}^{ij}}\delta^2 \tilde{X}^{ij}\right).$$
$$(32)$$

Here the first variation of $\tilde{X}^{ij}$ was already computed above and reads

$$\delta \tilde{X}^{ij} = \frac{1}{2}\tilde{\epsilon}^{\mu\nu\rho\sigma}F^{(i}_{\mu\nu}D_{A\rho}\delta A^{j)}_\sigma. \qquad (33)$$

The second variation reads

$$\delta^2 \tilde{X}^{ij} = \frac{1}{2}\tilde{\epsilon}^{\mu\nu\rho\sigma}D_{A\mu}\delta A^i_\nu D_{A\rho}\delta A^j_\sigma + \frac{1}{2}\tilde{\epsilon}^{\mu\nu\rho\sigma}F^{(i}_{\mu\nu}\epsilon^{j)kl}\delta A^k_\rho \delta A^l_\sigma.$$
$$(34)$$

III. CONSTANT CURVATURE BACKGROUND

In this and the next section, to get a better feel for our theory and also to prepare for its quantization, we consider the action (1) expanded around a specific background connection A^i.

A. Second-order action around a general background

We now write our connection as the background A^i plus a fluctuation $\mathcal{A}^i$, and obtain the part of the action quadratic in $\mathcal{A}^i$ directly from (32). Thus, we divide the second variation by 2, replace δA^i_μ by $\mathcal{A}^i_\mu$, and get the following Lagrangian

$$(8i) \mathcal{L}_{\mathcal{A}} = \frac{\partial^2 f}{\partial \tilde{X}^{ij} \partial \tilde{X}^{kl}} (\tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\rho} \mathcal{A}^j_\sigma) (\tilde{\epsilon}^{\alpha\beta\gamma\delta} F^k_{\alpha\beta} D_{A\gamma} \mathcal{A}^l_\delta)$$

$$+ 2 \frac{\partial f}{\partial \tilde{X}^{ij}} \tilde{\epsilon}^{\mu\nu\rho\sigma} (D_{A\mu} \mathcal{A}^i_\nu D_{A\rho} \mathcal{A}^j_\sigma + F^i_{\mu\nu} \epsilon^{jkl} \mathcal{A}^k_\rho \mathcal{A}^l_\sigma).$$

$$(35)$$

In the following works this action will be used for a background field method one-loop computation, but here we specialize to a particular background.

B. The background

The background that we take is a constant curvature one and can be defined as follows. As we have already mentioned, any connection defines a (conformal) metric, obtained by requiring the triple of curvature two-forms F^i to be (anti-)self-dual. Because of this, in practice, to specify a background it is easier to start with the corresponding metric, and then construct the connection, so that the triple of curvature two-forms for this connection is (anti-)self-dual with respect to the metric one started from. This is the procedure we follow.

So, we first describe the corresponding metric, and then use it to construct the background connection in question. Thus, let ds^2 be the interval for a constant curvature metric in four spacetime dimensions (de Sitter space). For our purposes it is convenient to describe it using the flat slicing so that the metric reads

$$ds^2 = a^2(\eta) \left(-d\eta^2 + \sum_{i=1}^3 (dx^i)^2 \right), \quad (36)$$

where η is the conformal time and x^i are the spatial coordinates. For the de Sitter metric the function $a^2(\eta)$ is a specific one, see below. The tetrad θ^I, $I = 0, 1, 2, 3$ associated with the above metric reads:

$$\theta^0 = a d\eta, \qquad \theta^i = a dx^i, \quad (37)$$

so that $ds^2 = \theta^I \otimes \theta^J \eta_{IJ}$, where $\eta_{IJ} = \text{diag}(-1, 1, 1, 1)$. As is known to anyone with experience with the Plebanski formulation of general relativity [5], it is very convenient to define the following set of objects (two-forms):

$$\Sigma^i := i\theta^0 \wedge \theta^i - \frac{1}{2} \epsilon^{ijk} \theta^j \wedge \theta^k, \quad (38)$$

where, as before $i = 1, 2, 3$. Explicitly, for the metric (36) we have

$$\Sigma^i = a^2 \left(i d\eta \wedge dx^i - \frac{1}{2} \epsilon^{ijk} dx^j \wedge dx^k \right). \quad (39)$$

As is not hard to check, in general the two-forms (38) are anti-self-dual

$$\frac{i}{2} \epsilon_{\mu\nu}{}^{\rho\sigma} \Sigma^i_{\rho\sigma} = \Sigma^i_{\mu\nu} \quad (40)$$

with respect to the Hodge star operation on two-forms defined by the metric $ds^2 = \theta^I \otimes \theta^J \eta_{IJ}$. Here the object

$\epsilon_{\mu\nu}{}^{\rho\sigma}$ is obtained from the volume form $\epsilon_{\mu\nu\rho\sigma}$ by raising two of its indices using the metric, and in our conventions $\epsilon^{0123} = +1$. Thus, (39) are anti-self-dual with respect to the metric (36).

Let us now introduce our background connection. It is SU(2) connection A^i_0 such that the covariant exterior derivative of Σ^i given by (39) with respect to A^i_0 is zero. In other words,

$$0 = D_{A_0} \Sigma^i = d\Sigma^i + \epsilon^{ijk} A^j_0 \wedge \Sigma^k. \quad (41)$$

It is not hard to solve this equation for A^i_0 explicitly. We get

$$A^i_0 = i \mathcal{H} dx^i, \quad (42)$$

where

$$\mathcal{H} = \frac{a'}{a}, \quad (43)$$

and the prime denotes the derivative with respect to the conformal time. It is not hard to show that the connection (42) is just the (anti-) self-dual part of the spin connection compatible with the tetrad (37).

We have not yet used (imposed) the condition that the background (36) is constant curvature. This condition can be written as

$$F^i(A_0) = M_0^2 \Sigma^i, \quad (44)$$

where we have introduced a dimensionful parameter M_0, with dimensions of mass. The Eq. (44) states that the curvature of A^i is a constant M_0^2. For the connection (42) this gives two equations

$$\mathcal{H}' = \mathcal{H}^2 = a^2 M_0^2, \quad (45)$$

where the second equality is the familiar Friedman equation. Its solution can be written as $a = (M_0(\eta_{\max} - \eta))^{-1}$, where $\eta_{\max}$ is an integration constant. This means that the physical time t (obtained from $a d\eta = dt$) is determined by the relation $(M_0(\eta_{\max} - \eta))^{-1} = e^{M_0 t}$, where a convenient choice of the integration constant was made. Thus we have $a = e^{M_0 t}$ and therefore an exponentially expanding Universe. The discussed constant curvature metric (de Sitter space) is of course a solution of Einstein's theory. The quantity M_0 is then related to the cosmological constant Λ via $M_0^2 = \Lambda/3$. In the case of GR the cosmological constant is a parameter of the theory. In the case of our theories, however, there is no similar parameter in the Lagrangian, so we will not in general be able to identify M_0 with any Λ, as there is no such parameter in the theory. However, we shall see below that a certain analog of Λ can be defined for any of our theories by evaluating the action on the background (42).

We thus take the constant curvature connection

$$A^i_0 = i M_0 a dx^i \quad (46)$$

as the background for the perturbative expansion of (1). Note that, as far as the background is concerned, the flat

limit $M_0 \to 0$ can be taken without any difficulty. In this limit $a \to 1$ and $A_0^i \to 0$.

C. Action evaluated on the background

In GR the gravitational action evaluated on the de Sitter metric is proportional to the volume of the Universe. Indeed, the Einstein-Hilbert action for the signature $(-, +, +, +)$ reads

$$S_{\text{EH}}[g] = -\frac{1}{16\pi G} \int (R - 2\Lambda)\sqrt{-g}\,d^4x. \qquad (47)$$

On a constant curvature background (in four dimensions) $R = 4\Lambda$ and we get

$$S_{\text{EH}}^0 = -\frac{\Lambda}{8\pi G} \int \sqrt{-g}\,d^4x. \qquad (48)$$

Let us see what our action evaluated on (46) gives us. Using

$$\tilde{\epsilon}^{\mu\nu\rho\sigma}\theta_\mu^0\theta_\nu^i\theta_\rho^j\theta_\sigma^k = \sqrt{-g}\epsilon^{ijk}, \qquad (49)$$

where $\sqrt{-g}$ is the square root of the determinant of the metric $ds^2 = \theta^I \otimes \theta^J \eta_{IJ}$, we easily get

$$\Sigma^i \wedge \Sigma^j = -2i\sqrt{-g}\delta^{ij}d^4x, \qquad (50)$$

where Σ^i are the anti-self-dual forms (38). Thus, the matrix $\tilde{X}^{ij}$ at the background is equal to

$$\tilde{X}_0^{ij} = -2iM_0^4\sqrt{-g}\delta^{ij}, \qquad (51)$$

i.e., is proportional to the identity matrix. Thus, the value of the action (1) on the background is

$$S[A_0] = -2M_0^4 f_0 \int \sqrt{-g}\,d^4x, \qquad (52)$$

where $f_0 := f(\delta)$ is the value of the defining function at the identity matrix $X^{ij} = \delta^{ij}$.

Thus, for the function f that corresponds to GR we expect (48) to be equal to (52) and thus

$$2M_0^4 f_0 = \frac{\Lambda}{8\pi G}. \qquad (53)$$

As is shown in the Appendix for the defining function of GR $f_0 = 9/(16\pi G\Lambda)$, and so the relation (53) holds. For diffeomorphism-invariant gauge theories different from GR we have neither G nor Λ parameters. The only parameters of the theory are those (dimensionless) parameters arising by expanding the defining function. The sole dimensionful parameter comes in when the background curvature parameter M_0 is chosen. The relation (53) then shows that we have a natural analog of the ratio Λ/G present in our theory, and given by the product of the defining function evaluated at the identity matrix times M_0^4. However, there is yet no natural way to define either Λ or G. Indeed, we could choose to expand a theory with given f around background with any value of M_0, so there

is no reason for the identification $M_0^2 = \Lambda/3$ as in GR. Similarly, without analyzing how our gravitons interact it is impossible to determine any analog of the Newton's constant for our theory. We note, however, that since for GR we have $f_0 \sim 10^{120}$, we should expect that for the defining functions of interest the value of the defining function on the identity matrix is extremely large. This knowledge will be of help when we analyze the graviton self-interactions.

D. Linearized action

We first check that the constant curvature background (46) is a solution of (11) and then evaluate the second variation of the action (32) at the background.

The derivatives of (3) at the identity matrix are easily computed. Let us first write down the general expression for the first derivative. We omit the tilde from X for brevity (we can always pull out the density weight factor from the function f using the homogeneity). We have

$$\frac{\partial f}{\partial X^{ij}} = \delta^{ij}\chi(X) + \text{Tr}(X)\chi_1'(X)\left(\frac{2X^{ij}}{(\text{Tr}(X))^2} - \frac{2\text{Tr}(X^2)}{(\text{Tr}(X))^3}\delta^{ij}\right)$$
$$+ \text{Tr}(X)\chi_2'(X)\left(\frac{3(X^2)^{ij}}{(\text{Tr}(X))^3} - \frac{3\text{Tr}(X^3)}{(\text{Tr}(X))^4}\delta^{ij}\right), \qquad (54)$$

where $\chi_{1,2}'(X)$ are the derivatives of the function χ with respect to the first and second arguments, evaluated at X. It is easy to check that for $X_0^{ij} \sim \delta^{ij}$ the second and third terms on the right are zero, and we have:

$$\frac{\partial f}{\partial X^{ij}}\bigg|_{X_0} = \delta^{ij}\chi(X_0) = \frac{f_0}{3}\delta^{ij}. \qquad (55)$$

We note that this is M_0-independent. We remind the reader that the background value X_0 of matrix X is given by (51) above.

Let us now compute the matrix of second derivatives of the defining function. Since the expressions in brackets in (54) become zero when evaluated on X_0, the only way to get a nonzero result in the second-derivative is to act by a derivative on these expressions. We get

$$\frac{\partial^2 f}{\partial X^{ij}\partial X^{kl}}\bigg|_{X_0} = \frac{2(\chi_1'(X_0) + \chi_2'(X_0))}{\text{Tr}(X_0)}P^{ij|kl}, \qquad (56)$$

where

$$P^{ij|kl} := I^{ij|kl} - \frac{1}{3}\delta^{ij}\delta^{kl}, \quad I^{ij|kl} := \frac{1}{2}(\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}). \qquad (57)$$

We have introduced a special notation $P^{ij|kl}$ for the matrix that appeared in (56), as this is just the projector on the symmetric traceless part $P^{ij|kl}\delta_{ij} = 0$, and similarly for the contraction with δ_{kl}.

Having evaluated the derivatives of the defining function at the background, we are ready to specialize (35) for our constant curvature background (46). However, let us first check that our chosen background is indeed a solution of

field Eqs. (11). With the quantity $\chi(X_0)$ being a constant, the background $B_0^i \sim F^i(A_0)$, and thus the field Eqs. (11) are satisfied (in view of the Bianchi identity).

Let us now consider the second term in (32). Since the matrix of the first derivatives is proportional to the identity matrix (55) with a constant proportionality coefficient we need to consider the integral of $\delta_{ij}\delta^2 \tilde{X}^{ij}$ over the manifold. Let us see that this is a total derivative. We have

$$\int_M d^4x \delta_{ij}\delta^2 \tilde{X}^{ij} = \frac{1}{2}\int_M (D_A \delta A^i \wedge D_A \delta A^i + F^i(A) \wedge \epsilon^{ijk}\delta A^j \wedge \delta A^k), \quad (58)$$

where we wrote everything in terms of forms (our form convention is $F = (1/2)F_{\mu\nu}dx^\mu \wedge dx^\nu$). Integrating by parts in the first term (and neglecting the total derivative term), the first term becomes

$$\frac{1}{2}\int_M \delta A^i \wedge D_A D_A \delta A^i = \frac{1}{2}\int_M \delta A^i \wedge \epsilon^{ijk}F^j(A) \wedge \delta A^k, \quad (59)$$

which is minus the second term in (58), and so (58) is a total derivative.

We therefore only need to consider the first term in (32). Let us write this directly in terms of the two-forms Σ^i by substituting the expression (44) for the background curvature. Using the anti-self-duality (40) of Σ^i we have the following compact expression for the second variation

$$\delta^2 S|_{A_0} = -g_0 \int_M d^4x \sqrt{-g} P^{ij|kl}(\Sigma^{i\mu\nu}D_{A_0\mu}\delta A_\nu^j) \times (\Sigma^{k\rho\sigma}D_{A_0\rho}\delta A_\sigma^l), \quad (60)$$

where we have introduced a notation

$$g_0 := \frac{\chi_1'(X_0) + \chi_2'(X_0)}{3}. \quad (61)$$

Note that the factors of M_0 have cancelled from this result. The combination (61) of the first derivatives of the defining function plays an important role below. Thus, we shall see that the constant determining the strength of self-interactions of our gravitons will be built from M_0 and g_0. We note that for the case of GR the quantity g_0 is of the same order as f_0 (see Appendix), and thus is very large.

E. High-energy limit

For applications in quantum gravity one is mostly interested in the UV behavior of the theory. Thus, we are interested in its behavior at energies $E \gg M_0$. In this case we can neglect the fact that the background is curved, and consider an effective theory in Minkowski space. As we shall see in this subsection, this certainly works at the limearized level. At the level of interactions we shall face some puzzles (related to the fact that the GR action blows up in the limit $\Lambda \to 0$), to be discussed below.

At energies $E \gg M_0$ the terms in the covariant derivative containing the usual derivative become much larger than the terms containing the background connection (the latter being of the order M_0). Thus, in the high-energy limit we can replace the covariant derivatives with the ordinary ones, and neglect the fact that the background is curved. However, the field δA^i in the linearized action (60) is not canonically normalized, as there is a numerical constant g_0 in front of the action. Absorbing this constant into the linearized fields by rescaling we obtain the following action

$$S_{\mathrm{lin}}[a] = -\frac{1}{2}\int_M d^4x P^{ij|kl}(\Sigma^{i\mu\nu}\partial_\mu a_\nu^j)(\Sigma^{k\rho\sigma}\partial_\rho a_\sigma^l), \quad (62)$$

where the (rescaled) linearized field is now called $a_\mu^i = \sqrt{g_0}(\delta A_\mu^i)$, and we have divided the second variation of the action by 2 to get the correct linearized action. The two-forms $\Sigma_{\mu\nu}^i$ are now those corresponding to the Minkowski spacetime

$$\Sigma^i = \mathrm{i}dt \wedge dx^i - \frac{1}{2}\epsilon^{ijk}dx^j \wedge dx^k. \quad (63)$$

Thus, in the high-energy limit one effectively works in the Minkowski background, and the connection perturbation has been rescaled to have a canonically normalized kinetic term. The operation of absorbing g_0 into the connection field is not that innocuous, as g_0 blows up in the limit $\Lambda \to 0$. But if one is not taking this limit, just considers the connection perturbations changing on scales much smaller than the scale of the curvature, then it is natural to absorb the (very large) quantity g_0 into the connection field to make it canonically normalized. We shall now study the action (62) in some detail, to see that it does describe the usual Minkowski spacetime gravitons. After this we turn to interactions.

A quick note about dimensions of all the fields. As we have already mentioned, we take the connection to have the mass dimension one, as is appropriate for a field that can be combined into a derivative operator. Then the curvature has mass dimension two, the matrix X^{ij} has mass dimension four, the matrix of first derivatives of the defining function is dimensionless, and the matrix of second derivatives has dimension minus four. The two-forms Σ^i that are constructed from the dimensionless metric are dimensionless. The constant g_0 introduced in (61) is a sum of derivatives of a function of dimensionless arguments, and thus is dimensionless. Overall, we see that the mass dimension of the integrand in (62) is 4, as needed.

F. Symmetries

We have started from a diffeomorphism-invariant action (1) and linearized it around the constant curvature (and then zero curvature) background. We should check that the linearized action that we have obtained is still diffeomorphism-invariant. As before, the diffeomorphisms can be lifted to the SU(2) bundle as follows:

$$\delta_\eta A^i_\mu = \eta^\alpha F^i_{\mu\alpha}(A). \tag{64}$$

Here η^μ is the vector field (of mass dimension minus one)—generator of an infinitesimal diffeomorphism, and $F^i_{\mu\nu}(A)$ is the curvature of A^i_μ. It can be checked that the above formula is a diffeomorphism corrected by a gauge transformation. Replacing the background curvature by its value (44) we get the following formula for an infinitesimal variation

$$\delta_\eta a^i_\mu = \sqrt{g_0} M^2_0 \eta^\alpha \Sigma^i_{\mu\alpha}. \tag{65}$$

This suggests that we consider vector fields $\xi^\mu = \sqrt{g_0} M^2_0 \eta^\mu$ of mass dimension one that are finite in the limit $M_0 \to 0$. Thus, let us consider the following variations

$$\delta_\xi a^i_\mu = \xi^\alpha \Sigma^i_{\mu\alpha}, \tag{66}$$

which will play the role of an infinitesimal diffeomorphism for the theory (62).

Another set of transformations that we have to consider are gauge symmetries. An infinitesimal gauge transformation is given by

$$\delta_\phi a^i_\mu = \partial_\mu \phi^i. \tag{67}$$

Let us now verify that the linearized action is invariant under (66) and (67). For this we will need the following identity

$$\Sigma^{i\mu\nu}\Sigma^j_{\nu\rho} = -\delta^{ij}\eta^\mu_{\ \rho} + \epsilon^{ijk}\Sigma^{k\mu}_{\ \ \rho}, \tag{68}$$

which can be verified by a direct computation. Here $\eta^{\mu\nu}$ is the Minkowski metric. Let us first consider diffeomorphisms. Thus, consider the quantity

$$\Sigma^{i\mu\nu}\partial_\mu\delta_\xi a^j_\nu = \Sigma^{i\mu\nu}\partial_\mu\xi^\alpha\Sigma^j_{\nu\alpha}. \tag{69}$$

Using (68) we see that ij-symmetric part of this quantity is proportional to δ^{ij}. However this, when contracted with the projector in (62), gives zero. Thus, the invariance under infinitesimal changes of coordinates is established. The invariance under gauge transformations (67) follows by noting that the quantity $\Sigma^{i\mu\nu}$ is antisymmetric and therefore $\Sigma^{i\mu\nu}\partial_\mu\delta_\phi a^j_\nu = 0$.

Since our gauge theory action (62) is both diffeomorphism- and gauge-invariant we can already make a suspected count of the number of propagating DOF. Indeed, the configurational variable of the theory should be the spatial projection of the connection. This has $3 \times 3 = 9$ components. Subtracting four diffeomorphisms and three gauge DOF leaves us with two suspected propagating DOF. Let us confirm this count by the Hamiltonian analysis of the linearized theory. This will also help us to see the gravitons explicitly.

G. Hamiltonian analysis

In this subsection we give a more detailed demonstration of the spin-two nature of our theory given in the introduction.

To obtain the action in the Hamiltonian form let us expand the quantity that appears as the main building block of the linearized action (62). We have

$$\Sigma^{\mu\nu}_i\partial_\mu a^j_\nu = \mathrm{i}\partial_i a^j_0 - \mathrm{i}\dot{a}^j_i - \epsilon^{kl}_i\partial_k a^j_l. \tag{70}$$

Here we have identified the spatial a and internal i indices using e.g. the component $\delta^i_a := \Sigma^i_{0a}$ of the background two-form, and ∂_i are the partial derivatives with respect to spatial coordinates. We raise and lower spatial indices freely using δ^{ij} metric.

It is now easy to compute the conjugate momenta. Since the time derivatives that appear in the action are those of the spatial projection of the connection, it is clear that only these components can have nonzero momenta. However, since the projector is involved in (62), we see that only the symmetric tracefree part of a^j_i has nonzero momenta. These are

$$\pi^{ij} = P^{ij|kl}(\dot{a}_{kl} - \partial_k a_{0l} - \mathrm{i}\epsilon_{kmn}\partial_m a_{nl}). \tag{71}$$

We note that the action (62) does not at all depend on the trace part of the spatial connection a^j_i. However, there is a dependence on the antisymmetric (and of course symmetric) parts. Let us separate the trace, symmetric and antisymmetric parts of a^j_i and write

$$a_{ij} = a^s_{ij} + b\delta_{ij} + \epsilon_{ijk}c_k. \tag{72}$$

Here a^s_{ij} is the symmetric and tracefree part, and b, c_i parametrize the trace and antisymmetric parts, respectively. Let us now rewrite the expression for the momentum using this decomposition. We have

$$\pi^{ij} = \dot{a}^{sij} - \mathrm{i}\epsilon^{ikl}\partial_k a^{sj}_l + P^{ij|kl}\partial_k(\mathrm{i}c_l - a_{0l}). \tag{73}$$

We note that the second term here is automatically symmetric and tracefree. On the other hand, it is clear that the Lagrangian density in (62) is

$$\mathcal{L} = \frac{(\pi^{ij})^2}{2}. \tag{74}$$

We see that the Lagrangian (density) is independent of b. This has a simple interpretation. Indeed, computing the infinitesimal diffeomorphism action on the temporal and spatial projections of the connection we find

$$\delta a^i_0 = \mathrm{i}\xi^i, \qquad \delta_\xi a^i_j = -\mathrm{i}\xi^0\delta^i_j - \epsilon^i_{jk}\xi^k. \tag{75}$$

This, in particular, means that the trace part b of the matrix a^j_i is a pure gauge quantity that can be set to zero by a temporal diffeomorphism. We also see that the Lagrangian depends on the antisymmetric part of spatial and temporal components of the connection only in the combination

$ic_i - a_{0i}$. Indeed, it is easy to check that precisely this combination is invariant under spatial diffeomorphisms, as the antisymmetric component transforms as $\delta_\xi c_i = \xi_i$. Let us denote the invariant combination by ϕ_i. As we shall soon see, it will become a generator of infinitesimal gauge rotations in our theory. Thus, we finally rewrite the momentum as

$$\pi^{ij} = \dot{a}^{sij} - i\epsilon^{ikl}\partial_k a_l^{sj} + P^{ij|kl}\partial_k\phi_l, \qquad (76)$$

and compute the Hamiltonian density as $\mathcal{H} = \pi^{ij}\dot{a}^s_{ij} - \mathcal{L}$. We get

$$\mathcal{H} = \frac{(\pi^{ij})^2}{2} + i\pi^{ij}\epsilon_i^{kl}\partial_k a_{lj} + \phi_i\partial_j\pi^{ij}, \qquad (77)$$

where we have dropped the index s from a^s_{ij} for brevity. Thus, now all the dynamical fields appearing in the Hamiltonian are symmetric tracefree tensors. The quantity ϕ_i is the Lagrange multiplier, which serves as a generator of SU(2) rotations on the connection. Indeed, the Poisson bracket of the integrated last term with the connection gives

$$\delta_\phi a_{ij} = \partial_{(i}\phi_{j)}, \qquad (78)$$

which is just the (symmetrized) gauge transformation. To see the structure of the arising Hamiltonian it is convenient to fix the gauge and require the connection to be transverse

$$\partial^i a_{ij} = 0. \qquad (79)$$

The momentum is required to be transverse by the condition obtained varying the action with respect to the Lagrange multipliers ϕ_i. So, it is now clear that the reduced phase space of our linearized system is parametrized by two symmetric, tracefree and transverse matrices a_{ij} and π_{ij}. This corresponds to two propagating DOF.

Let us now see what the dynamics becomes. To unravel the structure of the arising expression for the (reduced) Hamiltonian let us further rewrite it as

$$\mathcal{H} = \frac{1}{2}(\pi^{ij} + i\epsilon^{ikl}\partial_k a_l^j)^2 + \frac{1}{2}(\partial_k a_{ij})^2. \qquad (80)$$

Up to this point no reality conditions for the fields were specified. We can now deduce the linearized theory reality conditions from the Hamiltonian (80). Indeed, declaring the symmetric tracefree transverse connection field a_{ij} to be real, and defining a new real momentum field

$$p^{ij} := \pi^{ij} + i\epsilon^{ikl}\partial_k a_l^j, \qquad p^{ij} \in \mathbb{R} \qquad (81)$$

we can rewrite the linearized Hamiltonian in an explicitly positive-definite form

$$\mathcal{H} = \frac{1}{2}(p^{ij})^2 + \frac{1}{2}(\partial_k a_{ij})^2. \qquad (82)$$

The field equations that follow are now the usual

$$\Box a_{ij} = 0, \qquad (83)$$

which is just the wave equation for the two components of the connection field a_{ij}. This is how gravitons are described by our gauge theory approach. We note that one can recognize in the analysis of this section the linearized version of the new Hamiltonian formulation of gravity [22]. In particular, the arising reality conditions for the phase space fields are the same as in this formulation. Thus, even though our starting point of a gauge theory is a bit unconventional, the linearized theory mimics constructions familiar from other formulations.

What is different about our linearized theory (62) from the more familiar treatment in [22] is that no diffeomorphism constraints are left in the final result. Instead, our linearized action is simply independent of certain components of the connection field, so the theory is formulated on a smaller configuration space to start with. In other words, in our pure connection approach to gravity the Hamiltonian and diffeomorphism constraints of GR that usually require so much attention are solved once and for all by projecting out certain components of the connection field. This fact about our formulation must be very important for practical applications. And indeed, we shall see below that e.g. the issue of the gauge fixing is considerably easier here than in the case of metric-based GR.

IV. PROPAGATOR

In this section we invert the quadratic form that we have obtained by expanding the theory around the Minkowski spacetime background. In doing this we must decide on the gauge fixing.

A. Gauge fixing

We have seen that the action (62) is invariant under both gauge and diffeomorphism transformations, but we have also seen above that this invariance is manifested very differently in the two cases. Thus, in the case of the gauge invariance the situation is completely standard in that some of the field components have zero momenta and are thus Lagrange multipliers — generators of gauge symmetries. In the case of diffeomorphisms the situation is very different—we have seen that the action is simply independent of some components of the field, exactly those components that can be freely changed by performing a diffeomorphism. Thus, while there is very little choice for dealing with the gauge rotations—we have to treat them in the usual way by fixing the gauge and thus making the unphysical components of the gauge field propagate—we will need a different procedure for dealing with those components of the connection that gets affected by diffeomorphisms.

A useful analogy here is as follows. Let us consider a theory of two scalar fields ϕ, ψ with the Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu(\phi - \psi))^2. \qquad (84)$$

It is clear that the Lagrangian is invariant under a simultaneous shift of both of the fields by some function. The way this is realized is that the Lagrangian is simply independent of a certain combination of the fields, namely, of $\phi + \psi$, being only a function of the combination $\phi - \psi$. A natural quantization strategy in this case is to introduce a new field $\phi - \psi$ and rewrite the Lagrangian in terms of the new field only. Then only this combination of the fields is a propagating field, while the other combination $\phi + \psi$ is a fiction.

In the case of the simple Lagrangian above it is very easy to see what the propagating field is. In our case (62) this is much harder. In particular, we will not be able to rewrite the full Lagrangian in a way that has only diffeomorphism-invariant combinations of the connection components appearing (see, however, below for an expression for the linearized Lagrangian that depends solely on the "physical" diffeomorphism-invariant components of the connection). However, an appropriate strategy is as follows. We can consider the quadratic form (62) as a form on the space of diffeomorphism-invariant classes of connections a_μ^i, i.e. connections related via

$$a_\mu^i \sim a_\mu^i + \xi^\nu \Sigma_{\mu\nu}^i. \qquad (85)$$

The quadratic form in (62) is degenerate on this space because there is still the usual gauge invariance to be taken care of. However, this gauge invariance can be dealt with in the usual way, by fixing the gauge. As we shall see below, it will be possible to find a gauge-fixing condition that is invariant under (85). After doing this we obtain a nondegenerate quadratic form on the space of diffeomorphism classes (85). It can be inverted, to obtain a propagator on the space of diffeomorphism classes of connections. As is standard for gauge fixing, this procedure will make the temporal and longitudinal components of the connection propagate (and will add ghosts that will offset the effect of making these components propagating). At the same time, the components of the connection that are identified in (85) will not be propagating, as the propagator will involve a projector on the space of diffeomorphism equivalence classes. This way of dealing with the gauge symmetries of our theory is very different from the case of the metric-based GR, but is quite natural given that the diffeomorphisms are realized in our theory quite differently.

Having explained the logic of our procedure it remains to find a gauge-fixing condition that is diffeomorphism-invariant. After some trial and error we found the following gauge-fixing condition to be useful:

$$\partial^\mu \Pi^{\mu i|\nu j} a_{\nu j} = \frac{2}{3} \partial^\mu \left(a_\mu^i + \frac{1}{2} \epsilon^{ijk} \Sigma_\mu^{k\nu} a_\nu^j \right) = 0, \qquad (86)$$

where

$$\Pi^{\mu i|\nu j} := \eta^{\mu\nu} \delta^{ij} + \frac{1}{3} \Sigma^{i\mu\rho} \Sigma_\rho^{j\nu}$$

$$= \frac{2}{3} \left(\eta^{\mu\nu} \delta^{ij} + \frac{1}{2} \epsilon^{ijk} \Sigma^{k\mu\nu} \right) \qquad (87)$$

is a projector operator whose meaning is to be clarified below. The projector property

$$\Pi^{\mu i|\nu j} \Pi_{\nu j}{}^{\rho k} = \Pi^{\mu i|\rho k} \qquad (88)$$

can be checked by an elementary computation. It is easy to see that our gauge-fixing condition is diffeomorphism-invariant. Indeed, consider

$$\xi^\nu \Sigma_{\mu\nu}^i + \frac{1}{2} \epsilon^{ijk} \Sigma_\mu^{k\nu} \xi^\rho \Sigma_{\nu\rho}^j. \qquad (89)$$

Using the algebra (68) of $\Sigma_{\mu\nu}^i$ matrices we see that the last term here equals

$$\frac{1}{2} \epsilon^{ijk} \epsilon^{kjl} \xi^\rho \Sigma_{\mu\rho}^l = -\xi^\nu \Sigma_{\mu\nu}^i. \qquad (90)$$

Thus, the quantity in (89) is zero, and the gauge-fixing condition (86) is diffeomorphism-invariant. It is also clear that as far as the gauge transformations are concerned the last term in (86) is inessential, for it is zero for any a_μ^i that is a pure gauge $a_\mu^i = \partial_\mu \phi^i$. Thus, (86) is the usual gauge theory gauge fixing condition, corrected by a term that is inessential as far as the behavior under the gauge transformations is concerned.

Let us now confirm that the projector $\Pi^{\mu i|\nu j}$ is just that on diffeomorphism equivalence classes of connections, and so it is natural to apply it before the usual gauge-fixing condition is imposed (to make this condition diffeomorphism-invariant). We compute the action of the projector on the connection a_ν^j decomposed as in the previous subsection

$$a_\nu^j = a_0^j (dt)_\nu + (a_{ij}^s + b\delta_{ij} + \epsilon_{ijk} c_k)(dx^i)_\nu. \qquad (91)$$

The result is

$$\Pi^{\mu i|\nu j} a_\nu^j = \frac{2}{3} \left(\delta^{ij} \left(\frac{\partial}{\partial t} \right)^\mu + \frac{i}{2} \epsilon^{ijk} \left(\frac{\partial}{\partial x^k} \right)^\mu \right) (a_0^j - i c^j)$$

$$+ a_{ij}^s \left(\frac{\partial}{\partial x^j} \right)^\mu. \qquad (92)$$

We note the the quantity b got projected out, and the projected connection only depends on the temporal and the antisymmetric spatial components of the connection in the combination $a_0^i - i c^i$, as expected from the previous section. Thus, the projector $\Pi^{\mu i|\nu j}$ is indeed just that on the diffeomorphism-invariant subspace, and selects the components $a_0^i - i c^i$, which play the role of the generators of the Gauss constraints, as well as a_{ij}^s, which are the two propagating DOF plus three longitudinal modes of the connection. As usual for a gauge theory we shall make the components generators of the Gauss constraints as well

as the longitudinal components of the connection propagating by adding a gauge-fixing term, and then offset their effects by adding ghosts.

The projector $\Pi^{\mu i|\nu j}$ can be somewhat demystified by explaining what is its spinorial analog. Readers not familiar with the Penrose's spinor language [23] for gravity can skip this paragraph. Using spinors one can express the connection a_μ^i as a certain rank-4 spinor. Indeed, the space-time index gets replaced by a pair AA' of an unprimed and primed spinor indices. The SU(2) index i gets replaced by a pair AB of two unprimed indices, which is moreover AB symmetric. Thus we get $a_{CC'}^{AB}$ as our linearized theory dynamical field. It is now not hard to show that the projector $\Pi^{\mu i|\nu j}$ is simply that on the component of this spinor that is completely symmetric in all its three unprimed indices. Thus, schematically, $(\Pi a)_{ABCC'} = a_{(ABC)C'}$, where the brackets denote symmetrization. The projected out part is $a_{BA'}^{AB}$, and is thus a mixed rank-2 spinor and carries precisely four components, as is appropriate for something that can be projected out by a diffeomorphism. As we shall note below, the linearized action (62) can be written very simply in terms of the field $a_{(ABC)C'}$.

We now add the gauge-fixing condition squared with some parameter to the Lagrangian. Thus, we consider the following gauge-fixed Lagrangian on the space of diffeomorphism equivalence classes of connections

$$\mathcal{L}_{\text{gf}} = -\frac{1}{2} P^{ij|kl} (\Sigma^{i\mu\nu} \partial_\mu a_\nu^j)(\Sigma^{k\rho\sigma} \partial_\rho a_\sigma^l) \\ - \frac{\alpha}{2} \left(\partial^\mu a_\mu^i - \frac{1}{2} \epsilon^{ijk} \Sigma^{j\mu\nu} \partial_\mu a_\nu^k \right)^2, \quad (93)$$

where we have changed the order of indices jk in the gauge-fixing term for convenience, and absorbed the $(2/3)^2$ factor into the gauge-fixing parameter α. As in the case of Yang-Mills theory, the idea is now to select the gauge-fixing parameter α so that the gauge-fixed action is as simple as possible.

B. The algebra of gauge fixing

In this subsection we will simplify the expression for the gauge-fixed Lagrangian and find a useful value for the gauge-fixing parameter α. To this end, let us first write the Lagrangian in the momentum space. Omitting the argument $\pm k$ from the Fourier components $a_\mu^i(k)$ of a_μ^i for brevity we have the following expression

$$\mathcal{L}_{\text{gf}} = -\frac{1}{2} P^{ij|kl} (\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{k\rho\sigma} k_\rho a_\sigma^l) \\ - \frac{\alpha}{2} \left(k^\mu a_\mu^i - \frac{1}{2} \epsilon^{ijk} \Sigma^{j\mu\nu} k_\mu a_\nu^k \right)^2. \quad (94)$$

Let us expand the last term. Introducing a compact notation $(ka^i) := k^\mu a_\mu^i$ and expanding the product of two ϵ's we have

$$\left((ka^i) - \frac{1}{2} \epsilon^{ijk} \Sigma^{j\mu\nu} k_\mu a_\nu^k \right)^2$$

$$= (ka^i)^2 - (ka^i) \epsilon^{ijk} \Sigma^{j\mu\nu} k_\mu a_\nu^k + \frac{1}{4} \Sigma^{i\mu\nu} \Sigma^{i\rho\sigma} k_\mu k_\rho a_\nu^j a_\sigma^j$$

$$- \frac{1}{4} (\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{j\rho\sigma} k_\rho a_\sigma^i). \quad (95)$$

Let us now expand the first term of the Lagrangian. We have

$$P^{ij|kl} (\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{k\rho\sigma} k_\rho a_\sigma^l)$$

$$= \frac{1}{2} \Sigma^{i\mu\nu} \Sigma^{i\rho\sigma} k_\mu k_\rho a_\nu^j a_\sigma^j + \frac{1}{2} (\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{j\rho\sigma} k_\rho a_\sigma^i)$$

$$- \frac{1}{3} (\Sigma^{i\mu\nu} k_\mu a_\nu^i)(\Sigma^{j\rho\sigma} k_\rho a_\sigma^j). \quad (96)$$

We can now use the following two identities

$$\Sigma^{i\mu\nu} \Sigma^{i\rho\sigma} = \eta^{\mu\rho} \eta^{\nu\sigma} - \eta^{\nu\rho} \eta^{\mu\sigma} - i\epsilon^{\mu\nu\rho\sigma} \quad (97)$$

and

$$\Sigma^{i\mu\nu} \Sigma^{j\rho\sigma} - \Sigma^{j\mu\nu} \Sigma^{i\rho\sigma} = \epsilon^{ijk} (\Sigma^{k\mu\sigma} \eta^{\nu\rho} - \Sigma^{k\nu\sigma} \eta^{\mu\rho} \\ - \Sigma^{k\mu\rho} \eta^{\nu\sigma} + \Sigma^{k\nu\rho} \eta^{\mu\sigma}). \quad (98)$$

We can now use the identity (97) to rewrite the first term in (96), and the identity (98) to rewrite the last term as a multiple of the second plus some extra terms. We get

$$\frac{1}{2} (k^2 (a_\mu^i)^2 - (ka^i)^2) + \frac{1}{6} (\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{j\rho\sigma} k_\rho a_\sigma^i)$$

$$+ \frac{1}{3} (k^2 \epsilon^{ijk} \Sigma^{i\mu\nu} a_\mu^j a_\nu^k + 2(ka^i) \epsilon^{ijk} \Sigma^{j\mu\nu} k_\mu a_\nu^k). \quad (99)$$

We now note that if we make a choice

$$\alpha = \frac{2}{3} \quad (100)$$

then the terms $(\Sigma^{i\mu\nu} k_\mu a_\nu^j)(\Sigma^{j\rho\sigma} k_\rho a_\sigma^i)$, as well as $(ka^i)^2$ and $(ka^i) \epsilon^{ijk} \Sigma^{j\mu\nu} k_\mu a_\nu^k$ cancel out and we get the following simple gauge-fixed action

$$\mathcal{L}_{\text{gf}} = -\frac{k^2}{3} \left((a_\mu^i)^2 + \frac{1}{2} \epsilon^{ijk} \Sigma^{k\mu\nu} a_\mu^i a_\nu^j \right)$$

$$= -\frac{k^2}{2} \Pi^{\mu i|\nu j} a_{\mu i} a_{\nu j}, \quad (101)$$

where $\Pi^{\mu i|\nu j}$ is the projector (87). Because the projector on diffeomorphism equivalence classes appears here explicitly, it is obvious that this action is still invariant under the diffeomorphisms (85), and so is now a nondegenerate quadratic form on the space of diffeomorphism equivalence classes.

We note that the above analysis implies that our original linearized Lagrangian (62) can be rewritten (in the momentum space) in terms of the "projected" connection Πa schematically as follows:

$$\mathcal{L} = -\frac{1}{2}\left(k^2(\Pi a)^2 - \frac{3}{2}(k\Pi a)^2\right). \qquad (102)$$

Thus, our linearized Lagrangian is *different* from that for Yang-Mills theory for the projected connection Πa. Indeed, in the case of Yang-Mills the numerical coefficient in front of the second term in the brackets in (102) would be unity. In the case of Yang-Mills theory the value of the coefficient in front of $(ka)^2$ is fixed by the requirement of gauge invariance. The same is true in our case, and the different numeric value has to do with the fact that the projected connection Πa transforms under the gauge transformations in a more complicated way than $\delta a^i_\mu = \partial_\mu \phi^i$. Indeed, we have

$$\delta \Pi^{\mu i | \nu j} a^j_\nu = \Pi^{\mu i | \nu j} \delta a^j_\nu = \frac{2}{3}\partial_\mu \phi^i - \frac{1}{3}\epsilon^{ijk}\Sigma^{\nu j}_\mu \partial_\nu \phi^k. \quad (103)$$

It is this more involved transformation law for the projected connection that is responsible for the different from the Yang-Mills case numerical factor in front of the second term in (102).

We can now also note that our linearized Lagrangian in (62) admits a very simple description in terms of spinors. Thus, as we have already mentioned, in the spinor notation our connection a^i_μ gets described by a rank-4 spinor $a_{ABCC'}$. The diffeomorphism classes are described by the component which is symmetric in its three unprimed indices, or, in other words, by the $(3/2, 1/2)$ irreducible representation of the Lorentz group, where the first number denotes the representation in the space of unprimed spinors and the second one in the space of primed ones. The Lagrangian in (62) is then a multiple of

$$\mathcal{L} \sim (\partial^{(A}_{A'} a^{BCD)A'})^2, \qquad (104)$$

where the precise numerical coefficient is convention-dependent and will be spelled out elsewhere. Here $\partial_{AA'}$ is the two-component spinor Dirac operator. In words, the Dirac operator is used to convert the representation $(3/2, 1/2)$ described by the connection to the spin-two representation $(2, 0)$, and this is then squared to form the Lagrangian. The Lagrangian clearly only depends on the part $a_{(ABC)C'}$ of the connection, which makes it obvious that at least in the linearized theory the diffeomoprhisms are realized simply so that the action is independent of some of the connection components. The form (104) of the Lagrangian also explains the structure of the propagator that is obtained below.

C. Propagator

We now invert the quadratic form in (101). Thus, we add a current term to the action

$$S_{gf} = \int \frac{d^4k}{(2\pi)^4}\left[-\frac{k^2}{2}\Pi^{\mu i | \nu j} a_{\mu i}(-k)a_{\nu j}(k) + J^{\mu i}(-k)a^i_\mu(k)\right], \qquad (105)$$

and then integrate the field a^i_μ out. This can be easily done in the space of diffeomorphism equivalence classes, and we immediately see that the action with the original connection field integrated out is given by

$$S[J] = \int \frac{d^4k}{(2\pi)^4}\frac{1}{2k^2}\Pi^{\mu i | \nu j} J^i_\mu(-k)J^j_\nu(k). \qquad (106)$$

In other words, the propagator of our theory is given by

$$\langle a^{\mu i}(-k)a^{\nu j}(k)\rangle = (1/i)^2 \frac{\delta}{\delta J^i_\mu(-k)}\frac{\delta}{\delta J^j_\nu(k)}e^{iS[J]}\big|_{J=0}$$

$$= (1/i)\frac{1}{k^2}\Pi^{\mu i | \nu j}, \qquad (107)$$

which is just the usual $1/k^2$ term times the projector onto the space of diffeomorphism equivalence classes of connections, times the (convention-dependent) $1/i$ factor.

This finishes our discussion of the free theory of gravitons on the Minkowski spacetime background (or gravitons with energy $E \gg M_0$ much greater than the energy scale of our constant curvature background). We refrain from considering ghosts that are irrelevant for our purely classical purposes in this paper. Instead, let us now consider the lowest-order interactions.

V. INTERACTIONS

In this section we consider graviton self-interactions and discuss puzzles related to the fact that the action blows up in the $\Lambda \to 0$ limit.

A. Third variation of the action

The third variation of the action is easily computed from (32). We get

$$\delta^3 S[A] = (1/i)\int_M d^4x\left(\frac{\partial^3 f}{\partial \tilde{X}^{ij}\partial \tilde{X}^{kl}\partial \tilde{X}^{pq}}\delta \tilde{X}^{ij}\delta \tilde{X}^{kl}\delta \tilde{X}^{pq}\right.$$
$$\left. + 3\frac{\partial^2 f}{\partial \tilde{X}^{ij}\partial \tilde{X}^{kl}}\delta^2 \tilde{X}^{ij}\delta X^{kl} + \frac{\partial f}{\partial \tilde{X}^{ij}}\delta^3 X^{ij}\right). \quad (108)$$

We have already computed the first and second variations of the matrix $\tilde{X}^{ij}$ in (33) and (34). The third variation is given by

$$\delta^3 \tilde{X}^{ij} = \frac{3}{2}\tilde{\epsilon}^{\mu\nu\rho\sigma}D_{A\mu}\delta A^{(i}_\nu \epsilon^{j)kl}\delta A^k_\rho \delta A^l_\sigma. \qquad (109)$$

We also note that the fourth variation, of relevance for higher-order interaction vertices, is zero, which follows by expanding the product of two ϵ's and noting that there is always a δ^{ij}-contraction of two variations of the connection. On the other hand, spacetime indices of all four variations of the connection are contracted with $\tilde{\epsilon}^{\mu\nu\rho\sigma}$, and so the result is zero.

B. Cubic interaction

We have already computed the first and second derivatives of the defining function at the identity matrix in (55) and (56). Let us now compute the third derivative. Here we only consider a simpler case when the defining function depends on the invariant $\mathrm{Tr}(X^2)/(\mathrm{Tr})^2$. The general case will be described elsewhere. We get

$$\frac{\partial^3 f}{\partial X^{ij} \partial X^{kl} \partial X^{pq}}\bigg|_{X_0}$$
$$= -\frac{2g_0}{3(-2\mathrm{i}M_0^4)^2}(\delta^{ij}P^{kl|pq} + \delta^{kl}P^{ij|pq} + \delta^{pq}P^{ij|kl}), \quad (110)$$

where g_0 is the dimensionless constant given by (61), and $P^{ij|kl}$ is the projector on the symmetric traceless part that we already encountered above.

We now compute the cubic interaction term. Let us first discuss the simpler case when all three gravitons are high-energy $E \gg M_0$. In this case certain terms are dominant, and we are first going to describe these terms. We evaluate (108) at the constant curvature background connection (46). The last term in (108) is then seen to be a total derivative. We can also note that of the two terms coming from $\delta^2 X^{ij}$ one term is proportional to $(D\delta A)^2$, while the other is of the order $M_0^2(\delta A)^2$. Let us first neglect the term $M_0^2(\delta A)^2$ as compared to $(D\delta A)^2$. Then, after some rewriting we get

$$\delta^3 S|_{A_0} = \frac{g_0}{2M_0^2} \int d^4x \sqrt{-g}\, P^{ij|kl}(\Sigma^{i\mu\nu}D_{A_0\mu}\delta A_\nu^j)$$
$$\times [(\Sigma^{k\rho\sigma}D_{A_0\rho}\delta A_\sigma^l)(\Sigma^{m\alpha\beta}D_{A_0\alpha}\delta A_\beta^m)$$
$$- 3\mathrm{i}\epsilon^{\alpha\beta\gamma\delta}D_{A_0\alpha}\delta A_\beta^k D_{A_0\gamma}\delta A_\delta^l]. \quad (111)$$

Now passing to the high-energy limit $E \gg M_0$ we replace the covariant derivatives by the usual coordinate ones, and then rewrite the interaction term in terms of the connection field $a_\mu^i = \sqrt{g_0}(\delta A_\mu^i)$, for which the kinetic term (62) is canonically normalized. We also need to divide the third variation by 3! to get the correct (leading contribution to) the cubic interaction term. We get:

$$S^{(3)} = \frac{1}{12\sqrt{g_0}M_0^2} \int d^4x\, P^{ij|kl}(\Sigma^{i\mu\nu}\partial_\mu a_\nu^j)[(\Sigma^{k\rho\sigma}\partial_\rho a_\sigma^l)$$
$$\times (\Sigma^{m\alpha\beta}\partial_\alpha a_\beta^m) - 3\mathrm{i}\epsilon^{\alpha\beta\gamma\delta}\partial_\alpha a_\beta^k \partial_\gamma a_\delta^l]. \quad (112)$$

To summarize, schematically, the obtained leading-order contribution to the cubic interaction is of the form

$$\mathcal{L}^{(3)} \sim \frac{1}{\sqrt{g_0}M_0^2}(\partial a)^3. \quad (113)$$

We learn that our theory of gravity has a negative mass dimension coupling constant, and so is nonrenormalizable in the usual sense of the word, as could be expected. We also see that in our approach the self-coupling of our gravitons described by the connection perturbation a cannot be identified with the Newton's constant. Indeed, for the defining

function (A10) that corresponds to the cosmological constant GR we have $g_0 \sim M_p^2/M_0^2$. Thus, we see that the combination that appears in the denominator of (113), at least for the defining function that corresponds to GR, is given by

$$M_*^2 := \sqrt{g_0}M_0^2 \sim M_p M_0. \quad (114)$$

C. Discussion

Some remarks on the result (113) are in order. First, the obtained form of the cubic graviton self-interaction is different from that in GR. Indeed, the GR Lagrangian expanded (around the Minkowski metric) starts with the cubic interaction term $\kappa h(\partial h)^2$, where $\kappa \sim \sqrt{G} \sim 1/M_p$ and h is the metric perturbation. The GR cubic term is quadratic in the derivative operator, while (113) is cubic. This explains why the mass dimension of the coupling constant in (113) is minus two while in the cubic interaction term of GR it is minus one. Thus, both are nonrenormalizable by power counting, but the form of the interaction is different.

We also see that the coupling constant measuring the strength of self-interactions of gravitons in our approach is different from that in GR. This is not too surprising since in the usual metric-based approach the Newton's constant G sets the strength of interaction of gravitons with the stress energy of matter (or other gravitons). This is why it is a factor of $\sqrt{G}$ that serves as the theory's coupling constant. But the notion of the stress-energy tensor of gravitons is metric-based. Indeed, the stress energy arises as the variational derivative of the action with respect to the metric. In our approach gravitons are described in terms of a different field (connection a), and so the variation of the action with respect to a no longer has the meaning of the stress tensor. This is why the strength of self-interaction of the connection field a no longer needs to be directly identified with the Newton's constant.

This raises the question of how the Newton's constant can be identified in our theory. One way to do this could be to evaluate the four-graviton scattering amplitude, which in the usual metric-based approach is proportional to G. We leave this calculation to future work.

Another remark on (113) is that M_* given by (114) is the scale at which our perturbation theory appears to become strongly coupled. Thus, it appears that, unlike in the case of GR where the cutoff scale is M_p, the cutoff for the gravitational perturbation theory in the pure connection formulation is M_*. For the currently accepted value of the cosmological constant this is $M_* \sim 10^{-2}$ eV. Thus, it appears that our perturbation theory cannot be trusted for energies larger than 10^{-2} eV. While this fact would not be a problem for the envisaged renormalization group calculations that are nonperturbative in nature, this apparent strong coupling arising in our theory at such a low-energy scale should certainly be given an interpretation. This is, in

particular, worrying given the fact that for a particular defining function our theory is claimed to be the usual GR, with its very different strong coupling scale. There is clearly a puzzle here.

While we have not yet worked out a resolution of this puzzle in all details, we believe what happens is as follows. The first remark is that (113) is not the full cubic vertex, but only its part that blows up in the $M_0 \to 0$ limit. It is not hard to see that we have neglected another part (which goes to zero in the $M_0 \to 0$ limit) and that the full cubic interaction term is schematically

$$\mathcal{L}_{\text{full}}^{(3)} \sim \frac{1}{\sqrt{g_0} M_0^2} (\partial a)^3 + \frac{M_0}{\sqrt{g_0}} (\partial a) a^2. \qquad (115)$$

Thus, the cubic vertex consists of two parts. One blows up in the limit $M_0 \to 0$, which is not very surprising given that the action itself blows up in this limit (at least in the case of GR when we can identify M_0^2 with Λ). The other goes to zero in the same limit. The fact that there is a blowing up part seems to indicate that it is not possible to take the Minkowski spacetime limit, which would be very worrying given that we certainly would like to be able to scatter gravitons in Minkowski spacetime to be able to compare predictions of our theory to those of the usual metric-based formulation. However, it can be shown that the full interaction vertex (115) actually vanishes when all 3 external legs are put on shell. This is the same result as in GR, so in spite of some off shell blowing up terms the on shell result is completely the same as in GR.

Thus, to understand what happens one must consider higher-order interactions. One finds that the quartic interaction is schematically

$$\mathcal{L}^{(4)} \sim \frac{1}{g_1 M_0^4} (\partial a)^4 + \frac{1}{g_0 M_0^2} (\partial a)^2 a^2 + \frac{1}{g_0} a^4, \qquad (116)$$

where g_1 is a new coupling constant, related to higher derivatives of the defining function computed at the identity matrix. We see that there is again a blowing up leading-order term. The last term vanishes in the limit $\Lambda \to 0$, when $g_0 \to \infty$. However, we see that (in the case of GR) the second term is exactly the usual $(1/M_p^2)(\partial a)^2 a^2$ second-derivative graviton interaction. Thus, we see that when the four-graviton scattering amplitude is computed there is a blowing up contribution from the diagrams involving two cubic vertices, as well as another blowing up contribution from the quartic vertex. There are also finite contributions both from the quartic vertex as well as from the diagrams involving two cubic vertices. We believe that, when evaluated on the physical states, the blowing up contributions should cancel, while the finite pieces assemble into the usual GR result. We will not attempt such a calculation here as it requires technology (spinor helicity) that is beyond the scope of this paper. But the fact that the terms finite in the $\Lambda \to 0$ limit are

precisely of the familiar from GR two-derivative form support the picture sketched.

To summarize, the structure of interactions in our gauge-theoretic description of gravity is yet to be unravelled. It is, however, clear that the theory is as nonrenormalizable as the usual GR in the metric-based approach. What is different about our formulation is that the limit of the cosmological constant going to zero is a nontrivial one to take, for the action of the theory blows up in this limit. This is manifested in the fact that the interaction vertices contain blowing up pieces. Naively, this suggests strong coupling at a very low-energy scale. However, we believe that the issue is much more subtle and that when computed for the physical graviton states the scattering amplitudes are perfectly finite in the $\Lambda \to 0$ limit and for the case of the defining function corresponding to GR reproduce the known results. A verification of this is left to future work.

VI. CONCLUSIONS

In this paper we have proposed a new approach to the gravitational perturbation theory. While our main motivation was the quantum theory (renormalization), in the present paper we remained in the classical domain. We have recalled how a diffeomorphism-invariant gauge theory can be formulated using a homogeneous degree one defining function, and how such a theory for the gauge group SU(2) is a gravity theory describing two propagating degree of freedom. In particular, general relativity itself can be put in this framework, see the action (5). Our main interest here was in the perturbation theory. Hence, we expanded our general diffeomorphism-invariant gauge theory Lagrangian around a constant curvature connection (46). The original theory does not have any dimensionful parameters, and we have seen that it is the choice of the background that brings in a dimensionful quantity into the game, in our case the radius of curvature of the background. We then took a limit of the radius of curvature becoming very large (or working at energies such that the curvature of the background can be neglected). This way we obtained a theory on the Minkowski spacetime background.

The linearized action (62) we obtained is quite simple, and can be seen to be a natural construct involving the linearized connection, as well as the basic (anti-) self-dual two-forms $\Sigma_{\mu\nu}^i$. Indeed, as is sometimes done in the literature, one can introduce the derivative operators $\partial^{\mu i} := \Sigma^{\mu\nu i} \partial_\nu$. The basic building block of our linearized action is then $\partial^{\mu i} a_\mu^j$, where this quantity is symmetrized and then its tracefree part is squared to form the action. Note that the projector $P^{ij|kl}$ on the symmetric tracefree part is just that on the spin-two part of the tensor product of two spin-one representations, and this is another manifestation of how the spin-two appears in the game. Indeed, one could re-write our linearized gauge theory action using the spinor notation as a multiple of $(\Sigma^{\mu\nu(AB} \partial_\mu a_\nu^{CD)})^2$, where the

brackets denote the symmetrization. A completely symmetrized rank-4 spinor is the standard realization of the spin-two representation. Another, particularly clear way to rewrite our linearized Lagrangian is completely in terms of spinors, when all spacetime indices are eliminated in favor of the spinor ones. The Lagrangian then takes the extremely simple form (104). This should be compared to a much more involved linearized Lagrangian for gravitons in the usual metric-based approach. This considerable simplification of the linearized Lagrangian is in itself a significant plus of our approach.

Another very important feature of our approach is that diffeomorphism invariance can be dealt with in a very simple way. Recall that it is this gauge symmetry that is causing so much difficulty in any approach to gravity, perturbative or nonperturbative. In contrast, in our formulation diffeomorphisms can be dealt with once and for all, by simply projecting out certain components of the connection. We believe that this feature of our gauge-theoretic description of gravity is very important, to be fully appreciated with more work on this approach. In a certain sense, what replaced the usual diffeomorphisms in our approach are the SU(2) gauge rotations. We have seen that these must be gauge-fixed in the usual fashion. It is however much easier to deal with gauge rotations than with diffeomorphisms, something that can be appreciated from our derivation of the propagator of our theory. This propagator can be literally read off from the Lagrangian in its form (104). There is no such a simple derivation of the propagator for the metric-based gravitons.

We have also looked at the (cubic and quartic) graviton self-interactions as described in our gauge-theoretic framework. It was observed that the perturbation theory appears to become strongly coupled at a very low-energy scale $M_* = \sqrt{M_p M_\Lambda}$, and so appears to behave quite differently from the perturbation theory based on the Einstein-Hilbert Lagrangian. However, there are indications that this strong coupling may be only apparent, and that the physical scattering amplitudes are the same as in the metric-based GR. A verification that this is indeed the case is left to future work.

Apart from quantum aspects, which we purposefully decided to avoid here, we did not comment much on the subtle issue of the reality conditions for our theory. Indeed, these were discussed at the linearized level, where their treatment is no different from that in the Ashtekar formulation, see [22]. It is clear, however, that the full interacting action will require a much more sophisticated choice of the reality conditions. For the quantum calculations to be carried out with this formalism this is not much of an issue, because all loops are computed via the trick of the analytic continuation, and under this all factors of $\sqrt{-1}$ in our formulas disappear and fields become real. However, these issues do matter for the questions of the unitarity of the arising quantum theory. We expect that these subtle issues

will take some time to be settled, and refrain from trying to address them in this work.

To conclude, we hope to have convinced the reader that the present gauge-theoretic approach to gravity brings with itself many rather exciting opportunities that are simply unavailable, or impractical in the usual metric setting. It now seems within reach that, with the new tools developed here, the renormalization group flow for an infinite parametric class of gravity theories can be computed. Once this is achieved, ideas about the ultraviolet behavior of gravity, e.g. the asymptotic safety conjecture [18], can be explicitly tested.

APPENDIX A: DEFINING FUNCTION FOR GR WITH THE COSMOLOGICAL CONSTANT

In this appendix we derive an expression for the defining function corresponding to general relativity with a cosmological constant. We start with the Plebanski formulation [5] of the theory, and then integrate out the two-form field, as well as the Lagrange multiplier field. A similar in spirit derivation is given in [24]. However, the final result of that calculation is erroneous, see [25]. Here we present the correct defining function.

In the Plebanski formulation GR with a cosmological constant Λ is described by the following action

$$S[B, A, \Psi] = \frac{1}{8\pi iG} \int \left[B^i \wedge F^i - \frac{1}{2}\left(\Psi^{ij} + \frac{\Lambda}{3}\delta^{ij}\right)B^i \wedge B^j \right].$$
(A1)

Here G, Λ are the Newton's and cosmological constant, respectively, B^i is a $\mathfrak{su}(2)$-valued two-form field, A^i is a SU(2) connection, $i = \sqrt{-1}$, and Ψ^{ij} is the symmetric traceless field of Lagrange multipliers. More details on this formulation can be found in e.g. [6]. Integrating out the two-form field one gets the following action

$$S[A, \Psi] = \frac{1}{16\pi iG} \int \left(\Psi^{ij} + \frac{\Lambda}{3}\delta^{ij}\right)^{-1} F^i \wedge F^j, \quad \text{(A2)}$$

where it is assumed that the matrix $(\Psi^{ij} + (\Lambda/3)\delta^{ij})$ is invertible. It is now convenient to rescale the Lagrange multipliers field and write the action as

$$S[A, \tilde{\Psi}] = \frac{1}{i} \int (\tilde{\Psi}^{ij} + \alpha\delta^{ij})^{-1} F^i \wedge F^j, \quad \text{(A3)}$$

where

$$\alpha := \frac{16\pi G\Lambda}{3} \qquad (A4)$$

is a dimensionless quantity. Note that $\alpha \sim M_0^2/M_p^2$ and so is of the order $\alpha \sim 10^{-120}$.

In the final step we integrate out the Lagrange multiplier field $\tilde{\Psi}^{ij}$. Let us drop the tilde on the symbol for brevity. We can the rewrite the above action as

$$S[A, \Psi] = \frac{1}{i} \int (\text{vol}) \, \text{Tr}((\Psi + \alpha \text{Id})^{-1} X), \qquad (A5)$$

where we have introduced $F^i \wedge F^j = (\text{vol})X^{ij}$, and (vol) is an arbitrary auxiliary four-form on our manifold. To integrate out the matrix Ψ we have to solve the field equations for it, and then substitute the result back into the action. Assuming that the solution for Ψ can be written as a function of the matrix X that admits a representation as a series in powers of X, we see that Ψ will be diagonal if X is. Thus, we can simplify the problem of finding Ψ by using an SO(3) rotation to go to a basis in which X is diagonal. This is always possible at least locally. We then look for a solution in which Ψ is also diagonal. Denoting by $\lambda_1, \lambda_2, \lambda_3$ the eigenvalues of X^{ij}, and by a, b, $-(a + b)$ the components of the diagonal matrix Ψ, we get the following action functional to consider

$$F[a, b, \lambda] = \frac{\lambda_1}{\alpha + a} + \frac{\lambda_2}{\alpha + b} + \frac{\lambda_3}{\alpha - (a + b)}. \qquad (A6)$$

We now have to vary this with respect to a, b and substitute the solution back to obtain the defining function as a function of λ_i. Assuming that neither of the denominators in (A6) is zero we get the following two equations

$$(\alpha + a)^2 \lambda_3 = (\alpha - (a + b))^2 \lambda_1,$$
$$(\alpha + b)^2 \lambda_3 = (\alpha - (a + b))^2 \lambda_2. \qquad (A7)$$

Taking the (positive branch of the) square root and adding the results we get $a + b$, which is most conveniently written as

$$\alpha - (a + b) = 3\alpha \frac{\sqrt{\lambda_3}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}}. \qquad (A8)$$

The other two combinations that appear in (A6) are given by similar expressions:

$$\alpha + a = 3\alpha \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}},$$
$$\alpha + b = 3\alpha \frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}}. \qquad (A9)$$

It is now clear that the defining function is

$$f_{\text{GR}}(\lambda) = \frac{1}{3\alpha}(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})^2 = \frac{1}{3\alpha}(\text{Tr}\sqrt{X})^2. \quad (A10)$$

Thus, we learn that the action for GR with the cosmological constant can be rewritten in the form (1) as follows

$$S_{\text{GR}}[A] = \frac{1}{16\pi i G\Lambda} \int (\text{Tr}\sqrt{F^i \wedge F^j})^2. \qquad (A11)$$

For the defining function (A10) the value $f_0 = f(\delta)$ is given by

$$f_0 = \frac{3}{\alpha}, \qquad (A12)$$

which is thus of the order $f_0 \sim 10^{120}$. We can also compute the constant g_0. Thus, from (56) we get

$$\frac{\partial^2 f}{\partial \lambda_1 \partial \lambda_1} \bigg|_{\lambda_i = 1} = \frac{4g_0}{3}. \qquad (A13)$$

On the other hand, evaluating the second derivative of (A10) with respect to λ_1 we get

$$\frac{\partial^2 f_{\text{GR}}}{\partial \lambda_1 \partial \lambda_1} \bigg|_{\lambda_i = 1} = -\frac{1}{3\alpha}. \qquad (A14)$$

Thus,

$$g_0 = -\frac{1}{4\alpha}, \qquad (A15)$$

where the minus sign reflects the concave character of (B10). Thus we have $|g_0| \sim 10^{120}$ and $f_0/g_0 = -12$ for this defining function.

APPENDIX B: DEFINING FUNCTION FOR THE (MINIMALLY) MODIFIED GR

In this Appendix we analyze the defining function for what can be called minimally modified general relativity. The Plebanski-like action is given by:

$$S[B, A, \Psi] = \frac{1}{8\pi i G} \int \bigg[B^i \wedge F^i - \frac{1}{2}\bigg(\Psi^{ij} + \frac{\Lambda}{3}\delta^{ij} + \frac{\tilde{g}}{2}\text{Tr}(\Psi^2)\delta^{ij}\bigg)B^i \wedge B^j \bigg]. \qquad (B1)$$

Here $\tilde{g}$ is a constant of dimensions $\tilde{g} \sim M^{-2}$. Thus, this theory contains, in addition to G, Λ present in GR, an additional dimensionful coupling $\tilde{g}$. As before, we now integrate out the two-form field, and then rescale all the quantities by a multiple of $16\pi G$. Thus, we introduce:

$$g := \frac{\tilde{g}}{16\pi G}, \qquad (B2)$$

which is dimensionless, and then write the action omitting the tilde from over the symbol of Ψ. The resulting action is

$$S[A, \Psi] = \frac{1}{i} \int (\text{vol}) \text{Tr}((\Psi + \alpha \text{Id} + \frac{g}{2}\text{Tr}(\Psi^2)\text{Id})^{-1} X). \quad (B3)$$

Since it is natural to expect that the scale of deformation set by $\tilde{g}$ is of the order of $\tilde{g} \sim M_p^{-2}$, the natural values for g are order-1. The action then contains a small parameter α, and the action with Ψ integrated out can be found as an expansion in powers of this parameter.

As in the previous section we will integrate out Ψ by first diagonalizing X^{ij} and then looking for a solution for Ψ^{ij} as a function of X^{ij}, which guarantees that it is also diagonal. Thus, we have to consider the following functional of the eigenvalues only

$$F[a, b, \lambda] = \frac{\lambda_1}{\alpha + a + g(a^2 + b^2 + ab)}$$
$$+ \frac{\lambda_2}{\alpha + b + g(a^2 + b^2 + ab)}$$
$$+ \frac{\lambda_3}{\alpha - (a + b) + g(a^2 + b^2 + ab)}, \quad \text{(B4)}$$

where as before λ_1, λ_2, λ_3 are eigenvalues of X^{ij} and a, b, $-a - b$ are those of Ψ^{ij}. We now differentiate with respect to a, b and get the following two equations

$$\frac{\lambda_1(1 + g(2a + b))}{(\alpha + a + g(a^2 + b^2 + ab))^2}$$
$$= \frac{\lambda_3(1 - g(2a + b))}{(\alpha - a - b + g(a^2 + b^2 + ab))^2},$$
$$\frac{\lambda_2(1 + g(2b + a))}{(\alpha + b + g(a^2 + b^2 + ab))^2}$$
$$= \frac{\lambda_3(1 - g(2b + a))}{(\alpha - a - b + g(a^2 + b^2 + ab))^2}. \quad \text{(B5)}$$

We now look for the solutions in the form of a series:

$$a = a^{(1)} + a^{(2)} + \dots, \qquad b = b^{(1)} + b^{(2)} + \dots, \quad \text{(B6)}$$

where $a^{(1)}$, $b^{(1)}$ is $O(\alpha)$, $a^{(2)}$, $b^{(2)}$ is $O(\alpha^2)$, and the dots denote higher orders in the small parameter α. We have already found above that

$$a^{(1)} = \alpha \frac{2\sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}},$$
$$b^{(1)} = \alpha \frac{2\sqrt{\lambda_2} - \sqrt{\lambda_1} - \sqrt{\lambda_3}}{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}}. \quad \text{(B7)}$$

Using this we get

$$(a^{(1)})^2 + (b^{(1)})^2 + a^{(1)}b^{(1)}$$
$$= \frac{\alpha^2}{2(\sum_i \sqrt{\lambda_i})^2}(6\lambda_1 + 6\lambda_2 + 6\lambda_3 - 7\sqrt{\lambda_1\lambda_2}$$
$$- 7\sqrt{\lambda_1\lambda_3} - 7\sqrt{\lambda_2\lambda_3}). \quad \text{(B8)}$$

Let us introduce a compact notation

$$\Delta := 6\lambda_1 + 6\lambda_2 + 6\lambda_3 - 7\sqrt{\lambda_1\lambda_2} - 7\sqrt{\lambda_1\lambda_3} - 7\sqrt{\lambda_2\lambda_3}. \quad \text{(B9)}$$

Then functional (B4) can the be rewritten as follows:

$$F[a, b, \lambda] = \frac{\sum_i \sqrt{\lambda_i}}{3\alpha}\left(\sqrt{\lambda_1}\left(1 + \frac{(\sum_i \sqrt{\lambda_i})a^{(2)}}{3\alpha\sqrt{\lambda_1}} + \frac{\alpha g \Delta}{6(\sum_i \sqrt{\lambda_i})\sqrt{\lambda_1}} + O(\alpha^2)\right)^{-1} + \sqrt{\lambda_2}\left(1 + \frac{(\sum_i \sqrt{\lambda_i})b^{(2)}}{3\alpha\sqrt{\lambda_2}} + \frac{\alpha g \Delta}{6(\sum_i \sqrt{\lambda_i})\sqrt{\lambda_2}} + O(\alpha^2)\right)^{-1}$$
$$+ \sqrt{\lambda_3}\left(1 - \frac{(\sum_i \sqrt{\lambda_i})(a^{(2)} + b^{(2)})}{3\alpha\sqrt{\lambda_3}} + \frac{\alpha g \Delta}{6(\sum_i \sqrt{\lambda_i})\sqrt{\lambda_3}} + O(\alpha^2))^{-1}\right). \quad \text{(B10)}$$

Expanding the denominators in a power series in α and keeping only the $O(\alpha)$ terms we see that the terms involving $a^{(2)}$, $b^{(2)}$ cancel, and so we do not need to find these quantities to this order in α. We get the following functional

$$F[\lambda] = \frac{(\sum_i \sqrt{\lambda_i})^2}{3\alpha} - \frac{g\Delta}{6} + O(\alpha). \quad \text{(B11)}$$

We can rewrite it in a more convenient form by noting that the function

$$F_{\text{top}}[\lambda] = \sum_i \lambda_i \quad \text{(B12)}$$

gives rise to a total derivative, and so can always be added to our action. Thus, we can neglect multiples of $F_{\text{top}}[\lambda]$. It

is then easy to see that the function (B11) modulo (B12) is equal to

$$F[\lambda] \approx \frac{(\sum_i \sqrt{\lambda_i})^2}{3\alpha}\left(1 + \frac{7\alpha g}{4} + O((\alpha g)^2)\right), \quad \text{(B13)}$$

where $\approx$ stands for equal modulo $F_{\text{top}}[\lambda]$. The fact that it is the combination αg whose powers appear in brackets can be seen from (B4). Indeed, one can rescale the variables $a \to \alpha a$, $b \to \alpha b$ in (B4) so as to take $1/\alpha$ outside of the functional. Then the denominators will contain the combination αg, and it is clear that the function with a, b integrated out can be represented as an expansion in powers of αg. The first term in this expansion is given in (B13).

[1] R. P. Feynman, F. B. Morinigo, W. G. Wagner, and B. Hatfield, *Feynman Lectures on Gravitation* (Addison-Wesley, Reading, USA, 1995), p. 232.

[2] Z. Bern, Living Rev. Relativity **5**, 5 (2002).

[3] Z. Bern, T. Dennen, Y. T. Huang, and M. Kiermaier, Phys. Rev. D **82**, 065003 (2010).

[4] E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998).

[5] J. F. Plebanski, J. Math. Phys. (N.Y.) **18**, 2511 (1977).

[6] K. Krasnov, Gen. Relativ. Gravit. **43**, 1 (2010).

[7] A. Ashtekar, Phys. Rev. D **36**, 1587 (1987).

[8] T. Jacobson and L. Smolin, Classical Quantum Gravity **5**, 583 (1988).

[9] R. Capovilla, T. Jacobson, and J. Dell, Phys. Rev. Lett. **63**, 2325 (1989).

[10] I. Bengtsson, Phys. Lett. B **254**, 55 (1991).

[11] C. Rovelli, Living Rev. Relativity **11**, 5 (2008).

[12] K. Krasnov, arXiv:hep-th/0611182.

[13] I. Bengtsson, Mod. Phys. Lett. A **22**, 1643 (2007).

[14] K. Krasnov, Phys. Rev. D **81**, 084026 (2010).

[15] K. Krasnov, Europhys. Lett. **89**, 30002 (2010).

[16] A. Torres-Gomez and K. Krasnov, Phys. Rev. D **81**, 085003 (2010).

[17] K. Krasnov, Int. J. Mod. Phys. A **24**, 2776 (2009).

[18] S. Weinberg, Proc. Sci., CD09 (2009) 001.

[19] A. Torres-Gomez, K. Krasnov, and C. Scarinci, Phys. Rev. D **83**, 025023 (2011).

[20] K. Krasnov, Phys. Rev. Lett. **100**, 081102 (2008).

[21] J. Fine, A gauge theoretic approach to the anti-self-dual Einstein equations (unpublished) preprint.

[22] A. Ashtekar, *Lectures on Nonperturbative Canonical Gravity*, Advanced series in astrophysics and cosmology, 6 (World Scientific, Singapore, Singapore, 1991), p. 334.

[23] R. Penrose and W. Rindler, *Spinors and Space-Time. 1. Two Spinor Calculus And Relativistic Fields*, Cambridge Monographs On Mathematical Physics (Cambridge University Press, Cambridge, UK, 1984), 458.

[24] R. Capovilla, T. Jacobson, and J. Dell, Classical Quantum Gravity **8**, 59 (1991).

[25] R. Capovilla, T. Jacobson, and J. Dell, Classical Quantum Gravity **9**, 1839(E) (1992).